Defining and educating critical thinking

Pasquinelli, E., Farina, M., Bedel, A., Casati, R.

Report produced within the framework of Work Package 1
EEC Project - Critical Education (ANR-18-CE28-0018)

June 2020

This report is the result of the work of Work package 1 (WP1) of the EEC-Critical Thinking Education project (ANR-18-CE28-0018)

The objective of the present Report is to share the results of the EEC - Critical Thinking Education Project Work Package 1, dedicated to a. defining and characterizing critical thinking, b. identifying its specific cognitive underpinnings (naturalizing critical thinking), and c. deriving consequences for critical thinking education and transfert in daily life.

The EEC - Critical Thinking Education Project (ANR-18-CE28-0018)

In France, Critical thinking (CT) is a core aim of education, inesting disciplines such as science, history, language, and transversal teachings such as Media and Communication Education and Moral and Civil Education.

However, although educational projects are multiplying, there are still very few scientific studies evaluating the methods proposed by the different actors. Nor are these methods systematically based on existing scientific knowledge. The present project aims to fill this gap by designing and testing educational interventions for the development of critical thinking, thus establishing an evidence-based approach to critical thinking education. As a first step, the project aims to converge towards a more precise definition of critical thinking. It also aims to produce a scale for evaluating critical thinking. Finally, it proposes to disseminate these resources to institutions and teachers, as well as to the media and thes general public.

Objectives and work of Work Package 1 (WP1) - Theoretical analysis. Define and characterize critical thinking.

Coordinator: Roberto Casati (Institut Jean Nicod)

Authors: Elena Pasquinelli, Fondation *La main à la pâte*, Paris (France)**,** Mathieu Farina, Fondation *La main à la pâte*, Paris (France)**,** Audrey Bedel, Ecole Pratique des Hautes Etudes, Paris (France)**,** Roberto Casati, Institut Jean Nicod (CNRS/EHESS/ENS, DEC, Université PSL), Paris (France)

Thje following people have participated to the discussions of the WP1: Jie Gao (as part of a CogMaster PSL-P5-EHESS internship), Pierre Grialou (school teacher), Nathanaël Jeune (EdTech Master student, CRI), Sarah Peyronnet (school teacher), Calliste Scheibling-Sève (as part of her doctoral thesis), Emmanuel Trouche (Université Mohammed VI Polytechnique) and Sacha Yesilaltay (doctoral student, Institut Jean Nicod).

The following were listened to during the report preparation sessions: Grégoire Borst, director of the Lapsydé laboratory (University of Paris-Descartes), Hugo Mercier (CNRS, Institut Jean Nicod), Florent Meyniel (Neurospin Laboratory, CEA Saclay).
WP1 presented an interim report in June 2019 to the rest of the project consortium and received their comments and suggestions for improvement.

Characterization of the concept of "critical thinking" (CT): from existing approaches to a minimalist and operational definition

The main objective of WP1 was to provide a characterization of CT that clearly defines (1) a need and (2) an educational objective. To this end, the WP1 group analysed the literature on CT and its different definitions. Within this literature, the group sought to identify the elements to be retained for an operational characterization of CT. As the current approaches proved to be unsatisfactory from several points of view, the group turned to other approaches, outside the specific literature on CT, in order to fulfil the main objective.

Naturalizing critical thinking

WP1 identified a limited number of cognitive mechanisms and functions involved in the (critical) evaluation of information needed for decision-making and in confidence assessment. These functions constitute potential cognitive *building blocks* of CT and are therefore part of a cognitive approach to CT. These building blocks (1) serve as a basis for evaluating information (first and second hand) and (2) enable the confidence in the information thus evaluated to be adjusted in order to make a decision. The group made a special effort to identify minimalist functions and to develop an ecologically valid framework for CT. However, this goal remains at the draft level: the list of mechanisms and functions identified is not exhaustive, as their relationships have not yet been fully specified or brought together in a model.

Educability of CT

The WP1 group focused on the issue of capacity transfer and on the opportunity to embed the teaching of critical thinking in a disciplinary context. As soon as students demonstrate critical thinking in science, literature or mathematics, are they then able to transfer this skill to the learning of history or the evaluation of information? This issue is fundamental to CT education because the ultimate goal of such education can only be that of improving decision-making in everyday life, which presents very different contexts. The possibility of acquiring general skills independent of context and content is much debated. We have reviewed the literature on this issue and then focused on the more specific literature on how to transfer acquired capacities from one context to another, more or less distant one.
Finally, we listed several syntheses of the literature concerning the impact of current methods for CT education and the transferability of the results obtained. It appeared that the literature is extremely heterogeneous, which makes its synthesis difficult without taking the risk of comparing incomparable concepts. Among the methods that are supposed to be dedicated to CT education, some of them are indeed rather oriented towards text comprehension skills, towards the scientific method, others towards the recognition of biases, the acquisition of mathematical and scientific tools to counteract certain biases, etc. The methods used in the literature are not necessarily the same as those used in other fields.

Results of the study

The study has resulted in a minimalist, operational and cognitively realistic definition of CT, leading to principles to guide instructional practice and evaluation. In summary:
- CT is defined as the set of capabilities that allow one to assess the epistemic quality of the information available for decision-making, and to correctly calibrate one's confidence in this information according to the results of the assessment;

- the evaluation of CT thus consists in measuring the ability to put in place the most appropriate tools to assess the epistemic quality of information and to calibrate confidence accordingly.
- CT is not just an ideal to be achieved, but is part of the natural cognitive background of every individual, present from childhood onwards. We demonstrate CT when we select our sources based on reliability criteria or indices, when we judge the plausibility and relevance of information, when we feel or express confidence in the information available to us and in our choices or decisions. Children possess these abilities and use them often unconsciously. We consider these abilities to be part of *natural CT*;
- the natural abilities of CT are not flawless, but present structural limitations; e.g., the criteria or indices spontaneously used in order to assess the reliability of information and the trustworthiness of sources of information are not necessarily adapted to contemporary, complex, contexts and contents that require a more sophisticated CT;
- CT education aims to equip the natural CT capacities with increasingly sophisticated criteria. Its first objective is to put every citizen in a position to correctly evaluate the epistemic quality (validity, reliability) of commonly used sources of information and of common contents. We use the term "*advanced CT"* for this purpose. We distinguish between advanced CT and *expert CT*, which is applied to less common contents and which uses even more sophisticated criteria, specialized or even professional knowledge.
- In order to achieve its objectives, CT education is concerned with putting in place the best strategies to promote the transfer of skills and tools acquired in many fields of knowledge and contexts.

1. Introduction

1.1 Why is CT important?

Critical thinking is the flavour of the day, both in education and in public debate. CT education is invoked by teachers, parents, ministries of education and international organizations. Its supposed virtues range from promoting the capacity of students - as future citizens and workers - to apply "their knowledge in new and changing circumstances" (Howels 2018) and to enable relevant and credible information to be "selected, interpreted, digested, evaluated, learned and applied" (Halpern 2013). CT is now seen as more important than ever in meeting personal and societal challenges, in problem solving and in becoming more competent decision makers.

"Although the ability to think critically has always been important, it is a vital necessity for citizens in the 21st century" (Halpern 2013).

Among the circumstances that make the CT so relevant today, trends in the use of media and social networks play an important role: in the age of info-obesity, many point to the dangers of false news and post-truth attitudes, associated with a lack of basic skills to judge information correctly (Whitworth 2009; see Acerbi, 2019, for a critique of simplistic views of the impact of false news), while others stress the difficulty of protecting ourselves from our own biases (Pronin & Ross 2002). Simultaneously attacked from the outside and weak inside, we risk making bad decisions, supporting simplistic views that are not supported by evidence, and becoming gullible or, conversely, uniformly skeptical.

Many researchers working in the field of critical thinking deplore the poor state of critical thinking in most educated adults and children. For example, Halpern (1998) reports that many - if not most - adults do not think critically in many situations. She observes that many people have irrational beliefs, for example, about paranormal phenomena: they make irrational choices and do not look for relevant evidence. Kuhn (1999) describes the developmental pathway of metacognitive and epistemological skills that she considers crucial for CT - skills related to appreciating the role of evidence in accepting or updating theories and beliefs - to point out that these skills are almost as deficient in adults as in children.

Many others also conclude that most adults lack basic reasoning skills and are prey to inferences that "*frequently violate the principles of basic statistics, economics, logic and scientific methodology*" (Nisbett 2015).

Psychological traditions, such as the so-called "heuristics and biases" program, base their pessimism on the observation that individuals display a multitude of biases that violate the structure of rationality. *"For example, people display confirmation bias, test hypotheses ineffectively, display inconsistent preferences, do not calibrate belief levels correctly, project their own opinions too much onto others, combine probabilities inconsistently, and allow prior knowledge to be involved in deductive reasoning*" (Nisbett 2015; for summaries of the literature, see Evans 2012; Gilovich, Griffin & Kahneman 2002; Kahneman 2002; Shafir & LeBoeuf 2002; Stanovich 1999, 2009). The list of cognitive biases is long: base rate neglect, framing effects, representativeness bias, anchoring bias, availability bias, outcome bias, and liveliness effects, to name a few (Stanovich & Stanovich 2010). According to one view of human rationality, human beings are more model seekers and storytellers than they are naturally critical thinkers: "Indeed, *like ballet, critical thinking is a highly sought-after activity*. *Running is natural; nightclub dancing is less so; but ballet is something that people can only do well with many years of painful, expensive and dedicated training. Evolution didn't want us to walk to the end of our toes, and whatever Aristotle may say, we were not designed to be equally critical. "Evolution doesn't waste efforts to make things better than they should be...*" (van Gelder, 2005).

Proponents of CT education argue that the latter should be aimed at remedying this grim state of affairs, including shifting thinking patterns from the irrational to the rational: *"In short, much of the justification for educational interventions to change thinking dispositions stems from a tacit assumption that critical thinking dispositions make the individual a more rational person - or, as Sternberg (2001, 2005) argues, a wiser and less stupid person*. *Thus, the normative justification for stimulating critical thinking is that it is the foundation of rational thinking*" (Stanovich & Stanovich 2010).

1.2 But how?

How to "move from the irrational to the rational"? No simple answer exists yet for this question.

First, because the term "critical thinking" has several different meanings. This polysemy gives rise to a multiplicity of pedagogical interventions aimed at developing CT, but the contents of the interventions rarely overlap. What should be included in an educational programme aimed at strengthening critical thinking is, therefore, not yet clear. Existing methods are both heterogeneous and often very broad, encompassing a wide variety of skills and criteria. CT assessment tools suffer from the same problem. In addition, there are few tools that are intended to be adapted specifically for children.

Second, and key to our diagnosis, because most current approaches to CT do not address the natural cognitive building blocks that make CT possible in the first place. An effort is therefore required to build a theory of CT that can lead to operational outcomes (i.e., pedagogical methods and associated assessment tools). Such a theory should be based on a narrow (in the sense of specific) definition of CT. In this way, it would be possible to develop targeted actions and dedicated evaluation tools, and to compare the results of different pedagogical actions. We also postulate that a theory of CT should be informed by current knowledge in cognitive science, i.e., an understanding of the cognitive functions involved in CT, their developmental trajectories, and their limitations. The advantage of this approach is that it provides an objective ground on which pedagogical aims can be defined in terms of the development or equipping of pre-existing natural functions.

1.3 Methodology

We first review the approaches encountered in the CT literature. This work allows us to construct a new definition of CT, following the tradition of existing CT studies, but compatible with the operational objective.

Then, based on the proposed definition and through a review of the literature in developmental psychology, cognitive neuroscience, and evolutionary psychology, we identify a limited number of cognitive *building blocks* that make CT possible. We explore their developmental trajectories and highlight their limitations.

A better understanding of the natural foundations is a necessary condition for proposing scientifically based teaching strategies for the development of CT. We therefore propose a set of pedagogical principles that are compatible with the characterization of the natural capacities of CT, while also presenting their limits and room for improvement.

Particular attention is given to the issue of the transfer of CT-related competencies. This is a major concern for educators, especially since the possibility of educating so-called general skills such as CT has been repeatedly questioned.

2. Towards an operational definition of CT

2.1 Analysis of existing approaches. Observation of the absence of a consensus definition

In the academic literature, CT is primarily an **educational goal**, i.e. a desirable state that justifies a special educational effort (Hitchcock 2018, Lai 2012[1]). However, the term "critical thinking" has received and continues to receive a variety of definitions referring to different traditions, notably in philosophy and educational psychology, each with different views. **Some confusion thus persists as to the nature of critical thinking, despite the fact that the term has an increasingly important connotation.**

According to some authors, despite the apparent differences, existing definitions would converge around the same concept or conceptual "cluster": that of a **reflexive attitude**, allowing one to decide which beliefs to accept or refute; a **form of active and careful analysis of beliefs and the evidence in their favour**; a **disciplined process** that would allow one to analyze, synthesize, conceptualize, evaluate, and use information; a **mobilization of good reasons,** criteria; the **ability to self-correct; a** form of **open but also autonomous thinking**; an **inquisitive attitude** and **argumentative skills** (see, for example, Ennis 2016[2]).

Other authors (see, for example, Halpern 2013[3]) argue that at the heart of all definitions are **functions such** as reasoning, logic, judgment, metacognition, thinking, questioning, and other mental processes related to reasoning that lead to a solution or conclusion in a justified manner.

Still others recognize that **the consensus is about the goal of educating the critical mind rather than the definition of the term itself**. The lack of consensus would become particularly evident when moving from theory - and very general and inclusive definitions - to practice, which requires deciding on what to focus educational action (Bailin et al. 1999[4]). An example at this level is the definition produced, as part of the consensus-building effort, by a group of philosophers, social scientists, educators and scientists (physicists), under the leadership of the *American Philosophical Association,* in the late 1980s. To this end, the participants followed different structured stages of consensus-building, following the model known as the "Delphi process": at each stage, definitions are provided, commented on, and common ground is sought until a final and largely consensual definition of the capacities and dispositions for critical thinking is reached.

The definition that has emerged, however, applies only to an "ideal thinker":

"*We understand critical thinking to be the purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based. CT is essential as a tool of inquiry. As such, CT is a liberating force in education and a powerful resource in one's personal and civic life. While not synonymous with good thinking, CT is a pervasive and self-rectifying human phenomenon. The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit. Thus, educating good critical thinkers means working toward this ideal. It combines developing CT skills with nurturing those dispositions which consistently yield useful insights and which are the basis of a rational and democratic society.*" (Facione 1990[5]).

The authors consider that the effort to educate to improve critical thinking remains a worthwhile goal, and that the "ideal" thus defined should be used as a guide to envision CT education. However, the idealized nature of the definition makes it difficult to establish what is achievable and what is not.

Moreover, the definition is rather broad, in the sense that it refers to a wide variety of competences and provisions, which makes it even more difficult to imagine the transition to operationality. Nor are these skills and dispositions necessarily specific to critical thinking: they may, for example, be involved in effective communication, logical reasoning, problem solving. Indeed, the capacities listed as necessary for CT include capacities related to understanding a message, communication, expression, and justification, as well as others such as analysis and evaluation of arguments, inference, explanation, and self-regulation. In summary, CT comprises six dimensions or categories of capabilities: interpretation, analysis, evaluation, inference, explanation, and self-regulation. In addition to these capacities, the ideal critical thinker would possess dispositions to use these capacities and thus a "mind", inclinations, virtues. But, precisely, this is only an ideal: the panel could not reach a consensus on the advisability of including these provisions in the definition of CT.

Aware of this situation, educational psychologist Daniel Willingham - who is particularly interested in the pragmatic educational aspects of CT - has proposed a common sense definition of CT. According to this definition, CT is an effective form of thinking, i.e. one that leads to good solutions to a problem, that is voluntary, not automatic, and that does not rely on ready-made solutions already stored in memory (Willingham 2007, 2019[6]). However, this definition is, even according to Willingham, extremely general - it is equivalent to "good thinking" - and must in fact be applied specifically in each context and for each new content.

2.1.1 A bit of history. Philosophical approaches

Retracing, even summarily, the history of the different approaches to CT will allow us to better appreciate the similarities and differences between the various definitions, between the different existing CT measurement tools, and between the forms of intervention proposed to foster the development of CT.

Some trace the concept of CT back to antiquity, notably to Aristotle and the Sceptics. In this reconstruction, CT would be first and foremost a form of thought that follows the rules of logic. The "formation of the mind" would also be at the centre of Descartes' *Rules for the Direction of the Mind* and Hobbes' Theory of Reasoning; it would be found at the foundation of modern science in Francis Bacon and his critique of idols (*The Advancement of Learning),* in Robert Boyle and his model of skeptical thinking (*Sceptical Chymist)* (Paul, Elder, Bartell 1997; Foundation for critical thinking http://www.criticalthinking.org/pages/a-brief-history-of-the-idea-of-critical-thinking/4[7]). More generally, it is said to be a component of scientific thought that was asserted in the 17th and 18th centuries, the Age of Enlightenment.

However, it is only more recently, starting in the 1970s and 1980s, that CT training has become a theme of general reflection in education, notably thanks to a group of American philosophers who have taken up the idea of a strong link between critical thinking and education (Lai 2011; Lewis & Smith 1993). This tradition is strongly influenced by the seminal work of John Dewey, an educational philosopher who was very active in the 1930s.

Indeed, in the 1930s, CT became a pedagogical goal of schools that adhered to the Deweyen program (*Progressive Education Association*).

The founding role of John Dewey (1859-1952)

John Dewey had a strong interest in psychology and the study of the development of thought (Boisvert 2015; Dewey 1933, 1997). For him, school was an ideal place to study this development in a controlled environment, and to allow the child's natural abilities to develop in more varied and complex directions than those solicited by the natural, unguided experience. In Deweyen's educational model, the object of intellectual education is reflective thinking, which corresponds to critical thinking. Reflective thinking in fact coincides with the activity of evaluating the degree of probability that a belief is true, in order to accept or reject it, on the basis of an analysis of its foundations, and therefore of the facts that support it: *"Active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends, constitutes reflective thought."* (Dewey 1933[8])

Reflexive thinking involves two movements: on the one hand, an act of investigation and research: the study of facts and the examination and revision of evidence, the development of different hypotheses and their implications, the comparison of hypotheses with each other and with the results of observation and facts; on the other hand, the adoption of a sceptical attitude: to seek, one must first doubt, hesitate, be perplexed. Such an attitude is opposed to that which consists in accepting a belief on the basis of tradition, instruction, imitation - that is to say, of everything that has to do with authority - but also on the basis of the advantage that one hopes to derive from it, or of purely emotional reactions[9].

Dewey's definition does not focus on the outcome of CT education, on desirable behaviour, but on the processes involved - judiciously suspending judgment, determining the nature of the problem to be solved, evaluating arguments before making a decision or expressing an opinion (if necessary) - and is strongly linked to scientific reasoning. (To illustrate his definition, Dewey uses several examples and counter-examples[10]). According to him, humans are not naturally critical and the fact that we are intelligent organisms does not protect us from the spread of error and the accumulation of misconceptions. The credulity with which man is affected is dictated by the tendency to accept as true the first idea that comes to his mind.

Doubt (the first step towards critical thinking according to Dewey) has costs: it is painful and puts us in an uncomfortable condition of uncertainty. Moreover, not only human nature, but society itself, is not yet sufficiently rational, and so it instills bad habits of thinking through authority, instruction and imitation. However, the human child is just as naturally curious and spontaneously forms ideas and suggestions; he is, moreover, persevering. The education of critical thinking is therefore based on the exercise of these natural abilities and their transformation - through training of the mind - into habits of thought. This education constitutes the whole purpose of intellectual education. It must result in the production of a disciplined mind, which makes expert use of reflective thinking, that is, more specifically, a mind that knows:

(a) identify and define a difficulty ;

(b) develop a suggestion for its solution and reason about the consequences of the suggestion (using the capacities of deduction and induction);

(c) conduct further observations and experiments to accept or reject the suggestion (the formal stages of thinking).

However, it is also and above all a question of developing a mind that has a disposition to use these abilities, and this from a very young age. Science, especially the experimental method, uses all the operations of reflective thinking, and it does so in a systematic, elaborate and specialized way. The experimental scientific method is then the model to be adopted in education, to be applied to any problem, subject, belief, from primary school onwards. However, this does not mean that science education is particularly formative or should have a special place in instruction. On the contrary, the dimension of the scientific method must enter into all aspects of education: even manual, graphic, occupational disciplines, the most diverse activities, should and could be used to present students with problems to be solved through experimentation, and in such a way as to generate the need to acquire bodies of knowledge. This more specialized scientific knowledge is then the subject of specific education, in a second stage. Experimental, reflective and critical thinking thus becomes the aim of education, without constituting a separate teaching. In this sense, the training of critical thinking serves both science, as a knowledge enterprise par excellence, and the common good, because adopting one belief rather than another affects both our behaviour and other beliefs.

From the 1940s to the 1980s (from Edward Glaser to Robert Ennis to Matthew Lipman to the Center for Critical thinking to the Delphi report on Critical thinkling)

In the United States, the preoccupation with CT education continues in the 1940s with the work of Edward Glaser, an American educator. His influence is still perceptibel tosay, since Glaser's CT test is still in use, especially for the evaluation of CT in professional contexts (Glaser 1941[11], Watson 1980). For Glaser, the three pillars of critical thinking are :

- the reflexive attitude;
- knowledge of methods of logical or rational reasoning and investigation;
- the ability to implement these methods.

Attitudes and skills are therefore also seen as necessary in this tradition to foster CT.

In the 1960s, the interest for CT education is kept alive by reflections concerning the capacities that should be the subject of education. We may recall the taxonomy of cognitive abilities proposed by Bloom in 1965. This taxonomy places the acquisition of knowledge at the lowest level of his pyramid of abilities to be acquired. Next comes understanding, which requires going beyond knowledge, the application of what is understood, the analysis or critical evaluation of what is understood and applied, the synthesis of knowledge from different fields and, finally, the critical evaluation of knowledge that has been analysed and synthesised. The term "critical thinking" is not used, but the cognitive abilities listed are part of the higher-order group of abilities, which include critical thinking, problem solving, and metacognition, among others.

It is in 1962 that the philosopher Robert Ennis proposed a definition of the concept of CT as "*the correct assessment of statements*" (Ennis 1962), and then developed his definition to include a large number of capacities and dispositions that define the critical thinker (Lewis & Smith 1993; see,[12] for example, Ennis 1962, 1964, 1989, 1991, 2011, 2016). For Ennis, CT capabilities are divided into necessary and ancillary, but also general and specific (to a certain content). Necessary capabilities include:

- clarification, basic level: this includes focusing on the question we want to answer, analysing the arguments, asking or asking the questions necessary to clarify the problem;

- the evaluation of information, necessary to make a decision: judging the credibility of a source, based on criteria such as reputation, presence of justification, consistency, etc.; making observations or judging the quality of an observation, always on the basis of these criteria ;
- inference: logical deduction, generalisation, making value judgements based on facts or the consequences of the judgement ;
- advanced clarification: use of definitions, clarification based on pragmatic clues even about contents that are not explicitly expressed, use of assumptions ;
- making assumptions;
- synthesis.

These capabilities are listed as if they were applicable in any context. However, Ennis emphasizes both the importance of contextual adaptation and the importance of domain knowledge in combination with the listed capabilities. Ennis also gives attitudes and dispositions a fundamental place in CT. The ideal critical thinker seeks alternative explanations, is open-minded, willing to take seriously the views of others, willing to change one's mind in the face of new justifications, but also willing to go to great lengths to locate relevant information, to seek understanding, and to seek to clarify one's point of view in the face of others, he/she is an understanding, context-aware agent.

CT education therefore involves targeting both capacities and dispositions. The evaluation of the CT, on the other hand, involves examining the use of these same skills and the presence of the indicated dispositions (Ennis developed standardized tests, which are still widely used: Cornell's X and Z level critical thinking tests and the Ennis-Weir critical thinking test).

Almost in parallel with Robert Ennis's work, were developed the *Philosophy for Children* program and the *Center for Critical Thinking*.

The *Philosophy for Children* (P4C) program began in the 1970s in the United States at the initiative of Matthew Lipman (see, for example, Lipman 1987, 2003). Among its objectives, this programme includes the education of critical thinking skills. Lipman believed that children are capable of mobilizing CT, particularly when they are stimulated in this direction by the P4C program, and that they should be encouraged to do so from the time they enter school. His approach associated CT with the concept of judgement. Judgement is what allows opinions, evaluations, and conclusions to be formed. Thus, CT is the basis for making good judgements[13]. In this way, Lipman contrasts the ordinary judgment - the judgments we all make on a daily basis - with criteria-based judgments, the latter being specialized, advanced, and connected to CT ("*making judgment with criteria" = "good judgment"* and *"making judgment without criteria" = "ordinary judgment"*). As architecture is the professional way of looking at a building, CT is then the advanced way of exercising judgment.

"The outcomes of critical thinking are judgments; and the nature of judgment is such that critical thinking may be defined as skillful, responsible thinking that facilitates good judgment because (1) it relies upon criteria; (2) it is self-correcting, and (3) it is sensitive to context. The very meaning of 'criterion' is 'a rule or principle utilized in the making of judgments.' Judgment, in turn, is a skill; therefore critical thinking is skillful thinking, and skills can only be defined through criteria by which performance can be evaluated. So critical thinking is thinking that both employs criteria and can be assessed by appeal to criteria." (Lipman 1988[14])

Educating the CT therefore means providing access to these advanced judgement criteria, and CT education is what makes it possible to exercise our judgment in an advanced, "professional" way[15]. The notions of judgements and criteria are interdependent, yet, and CT education mostly relies on providing students with the right criteria as rules or principles used to make good judgements. Criteria are often specific to a certain area, hence the importance of paying attention to the context in choosing appropriate criteria and knowledge of the contents related to the context itself. In specific areas, these criteria are normally very well accepted and respected by experts in the investigative community, in that competent use of these respected rules is a means of establishing the objectivity of our normative, descriptive and evaluative judgements.

However, some criteria - such as consistency, reliability, and others - are meta-criteria because they have such a level of generality that they are always necessary for CT. This is why, in addition to the criteria, CT includes an advanced metacognitive component, which is not limited to looking at and monitoring oneself, but also involves self-correction.

Lipman also sees CT education as training for intellectual responsibility. By showing students models of epistemic responsibility, teachers invite students to take responsibility for their own thinking and, in a broader sense, for their own education.

Lipman's position is close to other philosophical approaches with respect to the use of judgment, criteria, metacognition, and of course the normative and voluntary nature of CT exercise. CT is a form of expert thinking that is equipped (by criteria), corrects itself, and takes into account the context. Contrary to other philosophical and psychological definitions we will encouter in the next paragraphs, Lipman's definition is rather restricted and operational, in that it consists of established criteria and a certain metacognitive attitude, and criteria can be identified and listed - even if, for the most part, they are domain specific.

Lipman does not, however, pose the problem of the day-to-day use of non-expert judgment: what natural cognitive bases allow it? What is used in alternative to advanced criteria? What are the pitfalls of natural judgment?

The 1980s are a sort of golden age for the spamming of interest in CT education. E.g., in the 1980s, the *Center for Critical Thinking* and the *Foundation for Critical Thinking were* created at the initiative of Richard Paul, philosopher, Linda Elder, teacher, and Gerald Nosich, philosopher (*Foundation for Critical Thinking*, https://www.criticalthinking.org/. Richard Paul is also the author of articles, a CT test, and several books containing methodological guidelines for his teaching (see: Scriven & Paul 1987, Paul & Elder 2009, Elder & Paul 2010). His approach is based on the dual observation of the universal nature of thought and the universally unsatisfactory quality of the products of thought. Indeed, human thought is, according to him, influenced by prejudices, mythologies, illusions, ignorance, and tends to be self-deceived. The capacities to be cultivated are therefore linked to the different processes that make up thought. Paul identifies eight of them: i. generating goals; ii. raising questions; iii. using the information ; iv. using concepts; v. making inferences; vi. making assumptions; vii. generating implications; viii. taking a point of view.

The critical thinker sets up these processes using certain capacities, like the non-critical thinker, but unlike the latter, the critical thinker adopts demanding intellectual standards for each capacity.

Paul emphasizes the importance of personality traits or dispositions alongside abilities: these are not sufficient to produce *strong-sense* CT but only *weak-sense* CT. These personality traits are also referred to as "epistemic virtues", which must therefore be cultivated in conjunction with the CE's own abilities in order to become a strong-sense critical thinker. Indeed, according to Paul, cultivating the CT abilities alone can lead to "doing good" as well as the opposite, because there is no guarantee that thinking abilities alone will lead to ethical choices. For CT to be used wisely for ethical purposes, it is therefore necessary to develop the epistemic virtues of intellectual empathy (which allows one to understand the point of view of others), integrity, perseverance, courage, autonomy, love of truth and reason, justice, and "*fair-mindedness*". Open-mindedness and intellectual honesty (*fair-mindedness*) is associated with CT in a strong sense. Paul thus distinguishes between a "Socratic" CT (which exercises these virtues in addition to abilities) and a "Sophist" CT (selfish, unethical), used to deceive others, to persuade them to abuse them. Paul also distinguishes between "explicit" CT and "implicit" CT (when abilities are used skillfully but without awareness of their role and how they can be exercised in a better way), "systematic" CT and "episodic" CT (the latter being exercised only on particular contents), "emancipatory" CT and "constrained" CT (which uses procedures and strategies but is not open to considering other points of view and alternatives), CT based on "natural" or "formal" languages (Elder et al. 1999), CT based on "natural" or "formal" languages (Elder & Co., 1999), and CT based on "natural" or "formal" languages (Elder et al. 1999, Paul 2010). In total, Paul and Elder (2009) list 35 "dimensions" of CT that include affective (intellectual virtues-related), macro-capacities, and micro-capacities.

In the end, CT is defined as a form of disciplined thinking that actively uses different intellectual capacities to guide belief formation and action. It is universal and general:

"*Critical thinking is the intellectually disciplined process of actively and skillfully conceptualizing, applying, analyzing, synthesizing, and/or evaluating information gathered from, or generated by, observation, experience, reflection, reasoning, or communication, as a guide to belief and action. In its exemplary form, it is based on universal intellectual values that transcend subject matter divisions: clarity, accuracy, precision, consistency, relevance, sound evidence, good reasons, depth, breadth, and fairness*." (Scriven & Paul 1987)

From a practical point of view, the teaching of CT is, for Paul, explicit and focused: it is necessary to insert the competences of CT in the curricula and to ensure that they are covered in all school subjects. A thorough learning of CT competencies requires long practice and instruction, but everyone can succeed. An essential condition for successful teaching of CT is that teachers themselves are able and willing to think critically, and that they have understood not only which competences they need to cultivate, but also the foundations of CT. In addition, it is important that they teach the skills of the CT, but also that they seek to develop dispositions for the CT (Elder & Paul 2010).

At the end of the 1980s, the growing interest in CT education led to a group of 46 philosophers and educators meeting as a Delphi panel between 1987 and 1989 to reach a consensus on the definition of CT, the capacities and dispositions involved, and to define modalities for measuring CT. The panel was coordinated by Peter Facione and commissioned by the *American Philosophical Association* (Facione 1990). The result is a normative definition of CT, broad enough to be agreed upon by the participants.

"CONSENSUS STATEMENT REGARDING CRITICAL THINKING AND THE IDEAL CRITICAL THINKER

We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based. CT is essential as a tool of inquiry. As such, CT is a liberating force in education and a powerful resource in one's personal and civic life. While not synonymous with good thinking, CT is a pervasive and self-rectifying human phenomenon. The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit. Thus, educating good critical thinkers means working toward this ideal. It combines developing CT skills with nurturing those dispositions which consistently yield useful insights and which are the basis of a rational and democratic society." (Facione 1990[16])

The experts of the Delphi report have also identified six fundamental capacities in which the good critical thinker should excel - but the bad critical thinker would have difficulty carrying out -: interpreting, analyzing, evaluating, making inferences, explaining, and self-regulating (Facione 2011[17]). However, these abilities are ideal.

"*The experts articulated an ideal. It may be that no person is fully adept at all the skills and sub-skills the experts found to be central to CT. It may be that no person has fully cultivated all the affective dispositions which characterize a good critical thinker. It may be that humans compartmentalize their lives in ways that CT is more active and evident in some areas than in* *others. This gives no more reason to abandon the effort to infuse CT into the educational system than that knowing no friendship is perfect gives one reason to despair of having friends. The experts' purpose in putting the ideal before the education community is that it should serve as a rich and worthy goal guiding CT assessment and curriculum development at all educational levels.*" (Facione 1990, Facione 2011[18]).

On the question of the need for the "good" critical thinker to also demonstrate moral virtues (using CT for good), the Delphi panel has not reached a consensual position: a minority considers that CT cannot be dissociated from the idea of ethics, while the majority considers CT as a set of provisions and capacities that do not involve moral (or political, religious, etc.) positioning.

Subsequently to the results of the Delphi Report, its leading figure, Peter Facione has developed CT measurement tools, such as the *California Critical Thinking Test* (CCTST), which draws directly on the definition that emerged from the work of the experts assembled for the Delphi Report. Facione has also developed tests specifically addressed to the provisions of the CT (CCTDI). Like other philosophers interested in CT, Facione questioned whether having the capacity to demonstrate CT was sufficient. He considers it unlikely that one would possess capabilities at an advanced level and not use them. However, the possibility of not using one's abilities does exist, and so a list of provisions must be provided that would make the critical thinker not only a skilled thinker, but also a kind of Sherlock Holmes motivated by the search for truth. (CT is indeed, according to Facione, the set of attitudes and abilities that give us the best chance of arriving at the truth.)

Thus, a critical thinker possesses, according to Facione (and in accordance with the vision of the Delphi panel) a certain attitude towards life - active, rational, open to the point of view of others, honest... - and shows certain attitudes when it comes to facing a problem, a question: clarity, systematicity, discipline, organization, attention, precision. Conversely, a weak critical thinker behaves in a disorganized and simplistic way when it comes to solving a problem, seeks the necessary information only in a punctual and non-systematic way, is easily distracted and is ready to drop everything when faced with the first difficulty, being satisfied with a vague and general answer (Facione 2011).

Facione refers to works in cognitive sciences concerning reasoning and in particular the articulation of thought in two complementary systems: intuitive and reflexive. He is very careful not to put them in opposition, and he considers them as two systems operating in parallel (and not sequentially). He thus refers to the notions of System 1 and System 2, of which he provides a simplified vision (indeed, the concepts of Type 1 and Type 2 thought processes are the subject of several definitions, and are subject to debate within the scientific community, as we will describe later in this Report). Teaching CT allows the development of the processes specific to System 2, but both systems are necessary for good decision-making. Thus, errors in decision-making and errors in judgement can arise - according to Facione's assertion - from both System 1 errors (heuristics that are misapplied or that wrongly influence a judgement) and System 2 errors (misrepresentation of facts, etc.). Facione goes on to describe a number of heuristics whose inappropriate activity is likely to lead us to error: availability, affect, association, simulation, similarity, satisfaction, risk and loss aversion, anchoring, illusion of control, retrospective bias, dominance (Facione 2011[19]).

From a practical point of view, in order to help the aspiring critical thinker develop or implement these functions, Facione proposes a *checklist* of questions to ask oneself in a situation that would require CT (this *checklist is* part of the CCTST manual for testing CT capabilities) and a point list for self-assessment of the readiness to demonstrate CT - or at least the dispositions that the subject would have demonstrated recently (this list is a simplified version of the CCTST or CT dispositions tests).

Elsewhere in America, Sharon Bailin and CT as inquiry

Sharon Bailin, a philosopher (Professor in the *Faculty of Education* at Simon Fraser University), represents a Canadian approach to CT that is different in many ways from the United States. E.g., Bailin and her colleagues deplore the fact that the notion of CT, although widespread, especially in curricula and curriculum discussions in the United States and the United Kingdom, remains vague and is the subject of only superficial consensus, which exists as long as practical questions about how to teach it, how to characterize the critical thinker, and what goals to achieve are not addressed. These authors recognize that the vagueness around the concept is a problem for educators (Bailin et al. 1999).

Bailin's approach to CT is fundamentally disciplinary: the goal of CT education is to "think well" within a discipline. Moreover, this approach is based on investigation. In fact, the term "inquiry" is used by Bailin and her collaborators (such as Mark Battersby, from the

Department of Philosophy at Capilano University) in a broad sense, and includes inquiry conducted in an argumentative manner, seeking the best arguments and reasons to support a position or opinion. CT teaching is thus integrated at the disciplinary level and is used to learn to use discipline-specific norms and modes of reasoning. In order to think critically and to learn to use norms and modes of reasoning, learning the concepts of the discipline is not sufficient. It is for this reason that disciplinary teaching needs to incorporate a specific approach to CT within it: reasoning and argumentation must receive explicit and focused attention in each disciplinary teaching. To this end, disciplinary teaching must integrate, on the one hand, aspects of argumentation that are specific to the discipline and, on the other hand, elements concerning argumentation that are of a more general, transdisciplinary nature. For example, it is important to become familiar with the different arguments that are debated within the discipline, historical debates, etc., in order to be able to understand the different aspects of argumentation. Without this disciplinary approach, it is not possible to learn how to reason and argue well within the discipline itself. For example, a course in ecology that integrates the CT dimension as an investigation would lead students to identify the context and the different arguments for and against a certain decision, and to ask themselves what are the rules for properly assessing the weight of each argument and its reasons.

Bailin acknowledges her affinity with the approaches of various philosophers and educators such as Richard Paul, Robert Ennis and Matthew Lipman. In the view of these thinkers, CT is a form of thinking that is directed toward a goal (as opposed, for example, to daydreaming) such as making a decision, answering a question, or solving a problem. More generally, it is a form of thinking that allows one to form a judgment, to develop an idea about what to believe or not to believe. Of course, it is not just any kind of thinking that leads to making a decision.
CT has a normative dimension: a decision made on the basis of poor reasons or a weak reasoning process would not count as CT. CT must therefore conform to standards or criteria for determining whether a reasoning process is good or poor. Furthermore, to be considered CT, these criteria and standards must be used consciously and voluntarily, not randomly, even though the person using them may not be able to formulate them verbally. According to Bailin, this type of definition is consistent with educators' understanding of CT (Bailin et al. 1999[20]).

However, Bailin points at the fact that existing definitions and approaches have major shortcomings, particularly in operational terms. First of all, they are often vague: neither judgments nor standards are defined. Second, the limits of CT in relation to other forms of thinking are unclear (e.g. what is the difference with problem solving, which requires judgements to be made in order to reach a decision). In addition, the concept of CT is sometimes tooe too broad and other times too narrow (as an example of the latter, Bailin cites Seigel's 1988 definition, which limits CT to the ability to properly assess reasons and the willingness and disposition to do so to guide one's choice of beliefs and actions. According to Bailin assessing the quality of arguments is not enough; one must also know which reasons are relevant, know how to take others into account, know how to see alternatives). Most importantly, for Bailin, current definitions do not take into account the fact that, most often, CT is required in a social context, in a discussion with an exchange of arguments. For this reason, Bailin proposes a rather different approach to CT education, based on critical discussion - what she calls investigation. The deployment of CT would therefore not be limited to the evaluation of reasons and arguments, but would include an exercise to properly respond to the arguments of others in the context of a discussion (Bailin et al. 1999[21]).

Once the social and argumentative dimension of CT has been established, it is necessary for Bailin to define the standards to which the evaluation of reasons and the exchange of arguments must conform in order to constitute a form of CT - recall that Bailin has a philosophical, normative approach: what matters is to identify, establish, the standards that lead to thinking and not to describe the capacities to think. To demonstrate CT would therefore consist in applying the standards that make reasoning correct. The capacities to be cultivated for CT education are those that make it possible to comply with the relevant standards[22].

Unlike the authors of the US tradition, Bailin refuses to use the terms "abilities" or "skills" to describe what the critical thinker must do. She thus refutes the seemingly descriptive and psychological jargon of other philosophers. She argues that what these philosophers actually indicate is a list of goals to be achieved, tasks that the critical thinker must be able to perform, not skills that the critical thinker would possess. Bailin thus resolves the ambiguity by choosing to call these "abilities" "results" that the critical thinker must aim for and achieve (Bailin et al. 1999[23]). Bailin is completely agnostic about the real cognitive abilities that need to be used and developed in order to achieve these goals. Furthermore, she believes that teaching CT does not mean teaching a set of abilities, but rather a set of resources. There are five types of resources (Bailin et al. 1999[24]):

- factual, domain-specific knowledge (*background knowledge*). The ability to think clearly, precisely and thoroughly in a field depends mainly on the knowledge one has in that field. The Bailin concept of CT is thus highly contextualized and dependent on context and content, in contrast to other more general concepts of CT. Indeed, Bailin makes CT dependent on the standards that are applied and on the use of those standards, and not on so-called cognitive abilities or skills. Even seemingly general abilities such as the ability to make a correct inference for a generalization depend on having a minimum amount of knowledge about the domain of the concept to be generalized;
- standards for critical appraisal. These standards, in turn, are domain-specific. For example, science has standards and norms that it applies to decide what is the best hypothesis, explanation, etc., for a given situation. For example, science has standards and norms that it applies to decide what is the best hypothesis, explanation, etc. To think critically in science, one must therefore know and know how to apply these standards, for example, the criteria that make a piece of data evidence of good or less good quality. However, the standards are not the same in science and art or in ethical choices. In order to think critically in each of these fields, it is therefore necessary to acquire the required standards. It is not a question of providing a list of them, or of learning them as theoretical elements: in order to be able to use them appropriately, it is necessary to understand the practices in which these standards operate. To apply the standards of science appropriately, for example, it is not enough to know the principles of the experimental method, but it is necessary to have a knowledge of science as practice. The standards of different disciplines or fields are cultural products, which evolve and can be subject to criticism;
- As each standard is domain-specific, the critical thinker must be able to recognize and distinguish between different domains in order to choose standards appropriately (e.g., distinguish between moral and scientific domains);
- the critical thinker also possesses heuristics. This term does not refer to natural cognitive heuristics, but to historically invented strategies for better understanding and

analyzing concepts, etc. The critical thinker also possesses heuristics. For example, the use of counter-examples is a heuristic that helps to clarify aspects of a concept;

- The critical thinker has habits of thought, attitudes, which put him in condition to use other resources. However, these habits are not only automatisms; the critical thinker has a critical mind: he recognizes the value of CT, its importance and, for this reason, he chooses to use the necessary resources. In this, Bailin agrees with the American tradition, which emphasizes dispositions, and she cites for this purpose the dispositions classically evoked by Paul, Ennis or the Delphi report, such as respect for reason and truth, questioning, willingness to investigate, *open-mindedness*, independence of mind, willingness to be fair (*fair-mindedness*), respect for others, respect for legitimate intellectual authorities, ethics of work well done.

If CT education aims to foster the use of concepts and standards that our culture has developed to improve thinking and make it more fruitful, the main way to achieve this goal, according to Bailin, is to become familiar, through examples, with the use of these principles, and to understand when and how they apply. Bailin therefore advocates contextualized teaching in rich, complex and realistic disciplinary contexts. This mode of teaching is opposed to presenting rules of thought and principles in an abstract manner. The teaching of CT must begin before school and continue on increasingly complex and specialized issues, with an increasing emphasis on the value of using these principles, and thus on a conscious and motivated use of CT. It is therefore a gradual process of learning (Bailin et al. 1999[25], Bailin & Battersby 2016). This process involves identifying the issue, identifying the context (identifying relevant contexts), understanding the competing cases, comparing and basing judgments on these comparisons (making a comparative judgment among them). This theoretical approach motivates a pedagogical choice consisting in putting learners in a real dialogical situation, i.e. presenting them with a discussion around a debated issue, providing the context and history of the debate. In the discussion, different arguments are presented for or against a certain position. Dialogues can deal with arguments as diverse as vegetarianism, the death penalty, the theory of evolution, violence in the media, human nature, polygamy, the interpretation of works of art, conspiracy theories... This type of intervention can therefore be inserted in any school subject. The pedagogical method also presents a theory of argumentation: an explanation of what constitutes a "good argument", the presentation of argumentative tricks, deduction and induction, analogical arguments, etc. In addition, it includes a methodology for the analysis of sources (internet sources, quotations, etc.). These elements constitute heuristics to be mobilized in order to make thinking more effective. Bailin's pedagogical method therefore focuses on how to carry out an investigation properly, i.e. how to identify the question being debated, how to differentiate the type of judgment one wants to give (evaluation or rather factual judgment). Learners then learn how to prepare their comparison of the arguments presented: they must first identify them one by one as being for or against the position being debated, identify the reasons given, etc. The learners then learn how to prepare their comparison of the arguments presented. Next, they must have a good understanding of the historical and social context of the issue being debated. Finally, they should analyse the strength of each argument made, using different criteria. If an argument is of an ethical nature, the criteria are of a different nature than if it is of a factual nature. It is at this point that it is possible to form an opinion based on the comparison of the arguments provided. However, the result is not guaranteed as if there were an algorithm to calculate the pros and cons in an objective way. Bailin points out that it is important to measure one's judgments against one's level of confidence. Finally, Bailin's pedagogical method proposes strategies for overcoming obstacles that make it difficult to keep an open mind in the face of other positions and arguments; indeed, Bailin acknowledges that simply urging people to remain open is not enough to maintain the "right attitude"[26]. Since Bailin's approach to CT is dialectical and argumentative, each chapter of her textbook (Bailin & Battersby 2016) begins with a contrasting narrative of critical and non-critical discussions on a topic. Readers are invited to pause and try to develop their own opinions on the issues presented (Bailin & Battersby 2016).

Bailin's pedagogical approach responds to a point she herself raised against the "general" methods of CT education: the nature of thinking is contextualized. This is not only for cognitive reasons - the fact that thinking is always that of an object in a context - but also for cultural considerations: we have norms and standards within our culture that allow us to think more expertly, clearly, precisely, in depth, and this within a variety of disciplines and with a different degree of development. If we want to learn to think better, we cannot detach ourselves from this cultural context. Then there is the problem of transfer. While Bailin's training of so-called capabilities, especially in the context of more generalist approaches, is unlikely to produce results, his approach may still make the CT so specialist that there is no hope of developing it in one area and then transferring it to another. Bailin is aware of this problem. Without learning the rules of a certain discipline or field, we cannot hope to succeed in thinking truly more expertly in that field. However, Bailin does not argue that there is no place for CT education upstream of the subject areas (in terms of school curricula). On the contrary, she advocates a progressive education that would begin before school entry. Her proposal, therefore, is to provide a variety of examples, to introduce CT teaching in all disciplines so that students learn a variety of standards and norms that apply to several areas. The student will thus become more effective in the type of judgment expected in different disciplines: physical sciences, social sciences, arts, etc., i.e., in a variety of contexts in which it is important for him or her to learn to think well. There are, however, problems that fall outside the classical school subjects, for example, moral problems or problems of daily life - as the example of vegetarianism shows - but even beyond: how to choose one's university? Who to vote for in the next elections? Bailin acknowledges that it is important to develop one's thinking habits and to acquire standards, knowledge and heuristics also in order to better consider these questions. She is therefore in favour of CT also being taught in specifically dedicated courses (and not only by infusion in the various disciplines), as long as it is taught in a dense context, and as long as the resources to be acquired are provided on the basis of concrete examples and rich in content and not as lists of standards, criteria or supposed capacities to be developed.

Common features of philosophical approaches

Philosophical approaches share certain points (Bailin 1999, Siegel 2010[27]):

- their nature is normative: philosophical definitions indicate the objectives to be achieved and the standards to be respected. From a theoretical as well as a practical point of view, it is a matter of identifying norms, standards, criteria to be applied in order to achieve a cognitive or practical objective: to make thinking more efficient in order to achieve a designated goal, to increase the chances of making a better decision, to embrace a correct opinion or one that is close to reality, to avoid developing false and/or unjustified beliefs ;

- These approaches are based on the polarization between thinking (thinking well) and not thinking: thinking is voluntary, explicit, and is the opposite of immediate action. They often refer to the idea of the 'ideal thinker': rational, motivated to seek the truth, and endowed with the capacity to do so;
- their nature is explicit: these definitions only refer to abilities used in a voluntary and assumed, reflexive manner ;
- These approaches most often provide a list of abilities attributed to critical thinking or critical thinking: metacognition in a very broad sense, judgment, rationality or reasoning and the use of reasons (giving and evaluating reasons or arguments), inference, logical inference, communication skills, expression, clarity, etc. This list is often long (even when a small number of categories are identified, they are then divided into sub-categories). The list is also very varied, and often not specific to the CT. For example, it includes skills such as the ability to interpret texts, to make inferences, to judge, to analyse arguments or to express oneself correctly;
- Critical point: the indicated abilities lack cognitive realism. Although the philosophical literature often speaks of "*skills*," "*abilities*," in reality, as Sharon Bailin points out, these abilities are not described in their "natural" functioning, but rather identified as goals to be achieved. These include, for example, the ability to evaluate information well or to make correct inferences. However, this literature does not describe how inferences are formed and therefore how to improve them, except by indicating what criteria to adopt when making an inference in order for it to be correct. The critical thinker is an ideal thinker who does not need to exist in reality because his or her characteristics are perfections of thought. Philosophical approaches often refer to the obstacles represented by natural thinking, but they do not ask whether the proposed perfections of thought (abilities but also dispositions) are realistic: under what circumstances can they be achieved? within what limits?
- Philosophical approaches refer, in addition to abilities, to dispositions, even epistemic virtues. However, the question of dispositions and virtues is still open in the philosophical debate around CT. Harvey Siegel places particular emphasis on the fact that the philosophical definition of CT has two components: one,

cognitive, related to capabilities, and the other, emotional, related to dispositions [he calls the first "reasoned assessment" and the second "critical spirit"], but that the two should not be considered separately. Thus, the idea of an ideal critical thinker takes on a strongly personal dimension, and becomes even more 'idealized' (Siegel 2010[28]);

- Representatives of the different philosophical approaches often propose CT teaching manuals and tests that tend to check whether the criteria, norms and standards are correctly applied, in a disciplinary or a-disciplinary way, in more or less realistic contexts, with a more or less remote transfer from learning situations. Some advocate a rather general teaching of CT, others a teaching infused in the disciplines. This depends on their vision of CT as a general or specific capacity. Several researchers consider domain knowledge (*background knowledge*) to be necessary for good thinking within the same domain (e.g. Bailin 2002; Willingham 2007). For others, general criteria exist that cut across knowledge domains, even if they have their own criteria (e.g. Lipman 1987; McPeck 1990). (For a discussion of this, see Ennis 1989, Lai 2011[29]).

2.1.2 Psychological approaches

In addition to philosophical approaches, there are a number of psychological approaches to thinking about CT. At the origin of these approaches we find a diverse group of psychologists such as Diane Halpern (cognitive psychology), Deanna Kuhn (educational psychology, developmental psychology), Richard Nisbett (social psychology, culture and cognition), Keith Stanovich (cognitive psychology, reasoning psychology), Robert Sternberg (psychology, psychometrics), Tim Van Gelder (philosophy), Daniel Willingham (educational psychology). Diane Halpern is the only one in this group to have proposed a fairly comprehensive textbook for teaching CT as well as a commercially available test (HTCA). Richard Nisbett has conducted studies on the effectiveness of teaching methods adopted by him and his colleagues with university students and adults. Keith Stanovich developed a theory of the functioning of the mind - the dual or "two systems" theory - which is often cited by other psychological but also philosophical (Facione) approaches as the basis for CT. Deanna Kuhn stands out because she has proposed the only developmental and truly descriptive approach to CT and has sought to identify the cognitive bases that would constitute the natural background for advanced, educated CT.

Diane Halpern and the cognitive approach to CT

Halpern sees CT as the set of capabilities that are necessary to succeed in complex tasks (Halpern 2013[30]). These tasks are quite varied and range from making decisions with societal issues to forming opinions and beliefs based on the best available information and solving problems. All CT capabilities are variable geometry due to context. However, these capacities are always directed in a voluntary and conscious manner (Halpern 1999[31]).

"*Critical thinking refers to the use of cognitive skills or strategies that increase the probability of a desirable outcome. Critical thinking is purposeful, reasoned, and goal-directed. It is the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions. Critical thinkers use these skills appropriately, without prompting, and usually with conscious intent, in a variety of settings. That is, they are predisposed to think critically."* (Halpern 1999)

These are therefore higher-order thinking skills - which, by definition, are necessary whenever we are faced with a complex problem - whereas using lower-order skills probably leads to a simplistic or wrong answer to the problem.

The specificity of CT lies however in its evaluative aspect. The concept of evaluation is broad and includes the evaluation of one's own reasons for acting in a certain way, the evaluation of the results of a certain thought process, the evaluation of the results of an action in problem solving, the evaluation of one's own decisions (Halpern 2013[32]). Halpern therefore contrasts CT (thinking, reflecting, learning concepts) and non-CE (no thinking, reflecting, rote learning), but also distinguishes in CT from other forms of thinking or high-level abilities (Halpern 2007[33], 1998[34]).

Halpern provides several taxonomies of CT-specific capabilities. These include the skills necessary, or at least useful, to increase the chances of solving a complex problem. The capabilities are divided into broad groups (macro capabilities): understanding of the problem and its context, questioning; using stored knowledge; skills related to language, analyzing the language used to carry arguments; inductive and deductive reasoning skills, logic in general; analyzing the structure of arguments; understanding probabilities; strategies and tools for problem solving; and creative thinking (Halpern 1999). The more synthetic taxonomy includes only five broad groups of abilities: verbal reasoning, argument analysis, hypothesis formulation and testing, probability reasoning, decision-making, and problem solving (Halpern 1998[35]).

Halpern's approach also has a normative dimension. It is not sufficient to exercise the cited capabilities to demonstrate CT. Certain criteria must be learned and applied, for example those that distinguish between arguments that are well supported by evidence (see Halpern 2013[36]). From a practical point of view, Halpern is optimistic about the possibility of teaching CE; she bases this on the results of large and small-scale experiments (Halpern 1999[37]). She is, however, aware of the difficulty of transferring skills to remote areas, as the acquired skills may not be activated by the new context. Transfer requires access to the deeper structure of the problem, beyond its superficial appearance (Halpern 1999[38]). From an operational point of view, Halpern therefore advocates explicit education of CT capabilities. However, this education would not be sufficient and should be accompanied by teaching the readiness to think and learn with effort; the teaching should be organized around the transfer of learned abilities and use explicit meta-cognition.

Halpern thus offers a teaching method with four components (Halpern 2013, Halpern 1999). (1) The first is learning to use the abilities that enable better thinking and to recognise situations in which these abilities are required (Halpern 2013[39]). But learning is not limited to developing these abilities or learning about them (for example, by reading a book about them). (2) In addition, one must be motivated to use them consciously, and be willing to check the truth of statements, to seek relevant information, to persevere when the solution to a problem is not immediate. Attitudes or dispositions that Halpern points out as underpinnings for CT are: a willingness to plan; flexibility in the sense of openness to consider many different and new ideas; perseverance; willingness to correct oneself, to change one's mind, to admit one's mistakes; an attitude of reflecting on oneself, on one's own procedures, rather than acting on autopilot; and a willingness to seek consensus within a group (Halpern 1999[40]). (3) EC learning also requires that the student be able and willing to monitor his or her thought processes, to check their effectiveness against objectives, and to decide when to allocate additional resources. For Halpern, metacognition is not a component of CT, but an accompanying function that facilitates the use of CT abilities. The instructor can help the student use metacognition by asking questions (Halpern 1999[41]). (4) Finally, the teacher should help the student transfer the general CT skills to different contexts and content. To do this, he or she must multiply the examples of application and make explicit the thinking skills required for CT (Halpern 1999[42]). In summary, the aim is to: (1) explicitly learning the abilities of CT; (2) developing dispositions for thinking that requires effort and learning; (3) learning in a way that facilitates transfer between domains; (4) making metacognitive monitoring explicit and conscious (Halpern 2013[43]).

Deanna Kuhn and the Developmental Approach to CT

Halpern's approach is cognitive - based on a set of cognitive functions or abilities - but does not take into account developmental factors. For example, in 1999, Kuhn noted that critical educational approaches - with the possible exception of John Dewey's approach, from which she drew inspiration - do not take sufficient account of empirical knowledge about the developmental trajectories of child and adolescent cognitive skills (Kuhn 1999[44]).

Like Dewey, Kuhn believes that the purpose of education is to become aware of the child's natural abilities and tendencies in order to develop them into more advanced attitudes and abilities. To do this, it is necessary to know the intellectual functioning of the child through empirical studies. In her studies of cognitive development, Kuhn has been particularly interested in how children manage to articulate theory - their prior knowledge - and data - new information, evidence - in their reasoning. The articulation between facts and theories is a fundamental aspect of professional science, and of knowledge construction more generally. Kuhn sees this articulation as a fundamental aspect of CT, defined as the ability to distinguish between knowledge and opinion, to decide what to believe, and to justify one's choice in a reasoned manner. According to Kuhn, this type of articulation requires the development of second-order, metacognitive processes that become the developmental cognitive basis of CT (Kuhn 1999[45]).

Kuhn specifically identifies three types of metacognitive processes involved in CT.

First, meta-knowledge, that is, reflection on the nature of beliefs and knowledge. The development of the ability to recognize the existence of false beliefs is crucial here, as it manifests the understanding that a belief is not a mere copy of reality. This development also provides an understanding that an assertion is different from the evidence that may support it. In turn, understanding the notion of evidence allows us to ask: "How do I know what I know?

What is the source of my knowledge?" It turns out that this type of ability follows a developmental trajectory. For example, the understanding of the distinction between facts and beliefs or facts and thoughts is still very limited in the three-year-old. Around the age of four, the ability to attribute false beliefs to others represents a milestone in this development, because the child is able - as many laboratory experiments show - to recognize that what he or she thinks may not correspond to reality. (It should be noted that, according to recent literature, the development of metacognitive abilities, particularly those related to false beliefs, is probably less linear and earlier, at least in an implicit form. See Beran, Perner, Proust 2012). The ability to positively link assertions and evidence develops later, and with more limitations, while showing individual variability. For example, preschoolers are often still unable to indicate, when they make an assertion, whether they are basing it on their own experience or someone else's story, or when and under what circumstances they acquired this knowledge. If one asks "how do you know?" they do not seem to distinguish between a factual finding and an explanation based on plausibility. For example, if preschoolers are shown vignettes of a race, showing a child wearing particularly "high-performance" shoes and, in one of the vignettes, that child lifting a trophy, and are asked "how do you know the child won? "The answers often do not distinguish between seeing the child lifting a trophy and the explanation "the child has shoes that make him or her run fast". The ability to make this distinction develops with age, starting at six years of age. The development of metacognitive abilities in children and adolescents is characterized by an increased ability to consciously and purposefully coordinate theory and empirical data, prior ideas and facts or evidence, as well as by the development of abilities to monitor strategies used to gain new knowledge (e.g., more effective strategies for memorization). Kuhn considers the development of these capacities to be fundamental to the progressive acquisition of scientific thinking. A critical point, however, is the observation that even an adult does not spontaneously obtain real mastery over these abilities. In other words, these capacities develop only imperfectly - at least in relation to an "ideal" represented in the form of scientific thinking deployed by professional scientists. During development, moreover, less successful strategies are not substituted by more effective and sophisticated ones. Rather, what changes is the frequency with which one is relied upon rather than the other, without the more primitive strategies disappearing altogether. Some limitations noted in preschool children persist in adolescents and adults (Kuhn 1999[46]).

The second metacognitive component of CT identified by Kuhn is strategic metacognition, which involves monitoring oneself and adopting more effective strategies for learning and remembering.

Finally, the third component is the epistemological one, which allows us to understand where knowledge comes from. Kuhn has a "constructivist" view of knowledge, according to which realism (knowledge comes from external reality) is the lowest level, from an epistemological point of view. Realism would ignore the complexities of the construction of knowledge, its partly subjective nature, and thus make it difficult to understand the disagreement between ideas and representations of the world. This is, according to Kuhn, the most important, or at least the most specific, component of CT. It is to understand that knowledge is neither a faithful copy of reality nor just a subjective opinion, but an articulation between ideas and facts, theories and data. This third component involves a vision of knowledge and science that is not devoid of a cultural component. Kuhn, however, envisages connecting it, like the others, to a developmental pattern (Kuhn 1999[47]). Indeed, during childhood and adolescence, the epistemic component evolves like the others, but without necessarily reaching a mature epistemological state: one that allows us to recognize both the objectivity and the influence of ideas on knowledge.

Educating CT maximizes the realization of intellectual potential. It is therefore a matter of fine-tuning a process whose final realization is already preconfigured in its early stages. The education of CT requires the clear identification of the developmental components of CT. This identification is all the more important if educators are to place their own CT education in a rich, disciplinary, content-dense context, as it is then even easier to lose sight of the pedagogical goal. The metacognitive perspective also helps to address the problem of dispositions: being willing to use one's abilities requires more than just the habit of doing so, it is about finding an intrinsic motivation, and reflecting on one's abilities and limitations can provide this motivation.

More recently, Kuhn has developed an argumentative theory of CT. In this second strand, Kuhn embraces a fairly general definition of CT as "*a form of reasoning that includes reflection, justification, application of reasoning*" (Kuhn 2018[48]). Arguing skills - analyzing arguments, counter-arguments, and understanding the value of different points of view - are then considered prerequisites for CT as defined. Kuhn has conducted several empirical studies to test the impact of the argument on CT. Her measures involve open-ended interviews to assess the ability to judge the credibility of a source, the structure of an argument and its quality, taking a position and considering opposing arguments, and metacognitive awareness (Kuhn 2018[49]). Kuhn therefore identifies these capacities with second-order or metacognitive capacities: the capacity to think about what constitutes knowledge, the capacity to think about how to obtain or justify knowledge. More generally, the ability to question and reflect on what it means to "know". Kuhn's interventions are quite rich, including metacognitive components, group work, discussion, and explicit instruction on how to analyze arguments. Even in a flawless experimental context (which is far from being the case for the study presented, which is why the results are not considered here), it would therefore be impossible to establish whether the argumentative aspects are the most effective aspects of the method, and whether, among these aspects, the most effective are those relating to direct instruction on how to analyse arguments or rather participation in discussion groups.

Richard Nisbett and *mindware* – a toolbox for advanced CT

A social psychologist known for his studies on fundamental error of attribution and cultural influences on thinking, Richard Nisbett does not offer a definition of CT. However, he is interested in teaching and improving certain thinking skills that are often considered to be CT, hence his place in this report. In addition, he has successfully implemented and tested "thinking skills" teaching. In addition, he is cited by Diane Halpern as an example of the possibility of educating "general abilities", hence CT.

Indeed, Nisbett defends a vision that runs counter to so-called general capacities: contrary to a long tradition going back to Thorndike (we will discuss this argument later in this Report), he believes that general capacities exist and can be successfully educated, success consisting in particular in the possibility of transferring what has been learned to distant contexts and contents, provided that the rules of resolution are the same. For example, Nisbett and his colleagues tested the effects of educating a group of subjects in the law of large numbers. Previous studies had already shown that this rule was generally understood - as an abstract rule, in the sense that it can be applied to a variety of contexts - by an average person. However, it was not adopted in certain contexts, particularly those that were less suitable for being considered quantitative. Both explicit instruction to the rule and implicit instruction *through* examples of application have had positive results on spontaneous application and transfer to contexts different from the original ones. This would confirm, according to Nisbett, the existence of a general rule (Nisbett et al. 1987[50]). But Nisbett's optimism about the educability of something like "reason" in general or CT is severely limited by the finding that there are not many general capacities that behave in this way. For example, rules of logical inference do not behave like statistical rules (Nisbett et al. 1987[51]). Nisbett draws the conclusion that only abstract rules such as statistical rules can be trained in a general way. It is therefore a question of identifying other general rules that behave in this way. Once put together, these constitute a formal education programme with applications in a variety of contexts.

In his book dedicated to equipping the human mind to make it more rational and efficient, Nisbett proposes a cognitive toolbox (*mindware*) composed of about a hundred concepts, principles, rules of inference developed by psychologists, statisticians, philosophers, logicians. These are tools that are in addition to common sense, natural intuitions (*supplements to* common-sense) and which are supposed to enable people to think and act more efficiently. Our common sense leads us, indeed, sometimes to make errors of judgment. Nisbett points out that most of our cognitive processes are silent inferences based on tacit knowledge. These inferences are the result of evolution, which has provided us with mechanisms to categorize, explain, argue, etc., in order to understand the meaning of our knowledge. The classic example is that of illusions, such as perspective illusions: our evolution took place in a three-dimensional world and inferring certain properties of objects from their position in space facilitates decision-making. However, two-dimensional drawings can mislead us. These considerations also apply to reasoning about categories (recognition and categorization of objects and individuals), about causes, etc. The result is a discrepancy between our "perceptions" of reality and reality itself (Nisbett 2015[52]). Nisbett mentions in particular the classical biases described by psychologists such as Daniel Kahnemann and Amos Tversky as *representativeness, framing, ignorance of chance and oversized perception of patterns, availability*. More generally, Nisbett lists the following causes of error in judgment:

- we are unaware that factors irrelevant to the problem silently influence our judgements (see the cited biases) ;
- we are ignorant of the context and give little consideration to its role in the behaviour of individuals (examples: fundamental error of attribution, effects of social influence);

- we are unaware that our choices are influenced by elements of which we do not have a conscious perception, we are unaware of the real causes of our opinions, choices, decisions;
- we are not aware of the importance of conducting a cost-benefit analysis before making a decision, of the influence of biases such as loss aversion;
- we do not know that many everyday situations can be described according to statistical rules of frequency, dispersion, mean and regression to the mean;
- we commit illusory errors of correlation and causality, again related to our ignorance of the statistical rules of association;
- we trust our assumptions, and therefore we do not test them scientifically (A/B test with or without randomization, with all the statistical considerations going with significance, covariance, etc.).

The principles and rules proposed by Nisbett are meant to act as corrective measures. According to him, they can be learned, become automatic and be used effortlessly in a variety of everyday problems (Nisbett 2015[53]). The solution proposed by Nisbett is essentially to become aware of the functioning of human cognition and of these biases and heuristics. This awareness should already make us humble (lower our confidence in ourselves and our judgements) (Nisbett 2015[54]). Thanks to this, it becomes possible to have recourse to cognitive tools that are more adapted to the situation. In order to be able to apply the right tools (rules, principles) at the right time, a subsequent effort must be made to :

- formulate and *frame* the problem so that the use of the rule becomes immediately relevant;
- codify the problem so that the principles can be applied to it. In the case of statistical rules, Nisbett's idea seems to be to familiarise as many people as possible with a statistical way of thinking applied to everyday life (Nisbett 2015[55]).

The approach proposed by Nisbett may raise several criticisms, in particular because the proposed teachings sometimes seem difficult to really import into everyday life. First of all, the need to code the problem at its "right level" of abstraction and to frame it in order to make salient the type of principle capable of solving it refers directly to the literature on transference and the difficulties of transference that Nisbett criticizes. The suggestion does not allow us to know, for example, which elements of context we should pay attention to, which elements silently influence us in a given context? Second, why, if these principles are so universal, are we so systematically mistaken? What is it that makes our cognitive functioning, though capable of adopting general rules, basically refractory to doing so - at least in the absence of dedicated education? Thus, when Nisbett suggests that we should perceive personality and behavioural issues as sampling problems, he neglects to explain why we do not do so spontaneously. What are we actually doing and why? In practice, Nisbett's approach does not seem sufficient to explain the discrepancy between our natural functioning and an augmented functioning that seems much more adaptive than the former. Thirdly, the idea of becoming aware of our functioning and our biases in order to become more humble is neither realistic nor necessarily desirable. Nisbett agrees with the idea of critical thinkers who seem to consider it positive that we systematically doubt ourselves, when this attitude would have disastrous consequences on our ability to act. Moreover, it is precisely contrary to the idea that evolution has provided us with ready-made solutions that are often useful and effective. In reality - without denying that our functioning, heuristics and biases sometimes mislead us - what we would often like to know is when a phenomenon is likely to occur, in order to learn how to identify and anticipate risk situations.

The notion of tools that make us smarter is a very interesting notion, however. According to Nisbett, we have become smarter over the course of our cultural evolution. We have invented mathematical and statistical tools, the scientific method, the notion of hypothesis, of test. These tools are taught systematically, but normally they are not taught in relation to the idea of developing CT and are not taught in a way that facilitates their use in everyday life. They should, in order to tool our CT skills.

Keith Stanovich: CT = rationality

Keith Stanovich is one of the leading representatives of the *dual system* models of thinking[56]. In these models our cognitive organization is postulated to fall into two main typologies of processes, sometimes called "systems": type 1 processes and type 2 processes, or System 1 (S1) and System 2 (S2). S1 is an operating modality (not a system in the sense of a particular cognitive architecture, a module) based on general and rapid heuristics of problem solving or judgement; S2 is an operating modality based on the implementation of *ad hoc* algorithms for the situation, chosen deliberately. According to Stanovich, the two systems operate in parallel: System 1 is always active, and System 2 intervenes or overrides System 1 on certain occasions. In reality, there are a variety of dualist models, and each model has undergone changes over time (Evans & Stanovich 2013). Stanovich even added a third system, S3, which arises from the decomposition of S2 into an analytical and a reflexive system. S3 comes into play when an algorithmic conflict between S1 and S2 is detected; its purpose is to resolve the conflict (Evans & Stanovich 2013).

The theory of CT proposed by Stanovich falls within this framework. For him, the study of CT necessarily has a normative dimension: its objective is to teach and learn to think better (Stanovich & Stanovich 2010[57]). But what does "better thinking" mean? And why should better thinking be positive? Stanovich solves the question by linking the idea of better thinking with the idea of rationality, rational thinking. Rational thinking is the one that is most likely to lead us to form beliefs that correspond to reality and allow us to achieve our goals. Abilities or attitudes such as open-mindedness are at the service of this rationality, but do not constitute its heart. Stanovich thus establishes a hierarchy between the capacities of CT, with a final objective to be achieved, which is that of rationality, and capacities, which may constitute means or associated capacities. If we consider openness and other capacities indicated in the literature as specific to CT, then we must consider CT education as having a higher "goal": development of rational thinking. More broadly, this means linking the concept of CT to that of rationality. According to Stanovich, this association between rationality and CT has several advantages, including anchoring the relatively vague concept of CT on a more developed concept in cognitive science research, both theoretical and empirical (Stanovich & Stanovich 2010).

In his definition of rationality (and thus of CT), Stanovich cites two components, one epistemic and one pragmatic or instrumental. Both are studied extensively in cognitive science[58]. On the one hand, epistemic rationality is defined in relation to the concepts of well-calibrated confidence and available evidence: the degree of confidence in an assertion is well-calibrated in relation to available and relevant evidence. On the other hand, pragmatic rationality is associated with the ability to optimally satisfy one's objectives. The two forms of rationality are linked, the first serving the second. Rationality would therefore allow us to maximize utility, through the possibility of following certain patterns of rational choice (Stanovich & Stanovich 2010). However, according to the theory of heuristics and bias, rationality is often violated on a daily basis because of numerous biases documented in the literature. For Stanovich, therefore, biases are harmful influences that lead us to move away from an "ideal" and ideally desirable rationality to the optimization of choices. Stanovich does not take into account the evolving reasons for bias or the fact that the heuristics in question most often have a positive impact on the decision, which they make optimal in a specific context (Stanovich & Stanovich 2010). Two important concepts in Stanovich's CT/rationality theory are therefore those of prior beliefs and bias. For Stanovich, the essential characteristic of CT is the ability to evaluate evidence and arguments independently of our own prior beliefs, opinions, and any biases or associations that lead us to give automatic, unconscious, non-reflexive responses in relation to context. In other words, it is unbiased thinking, capable of de-contextualization. Stanovich points out that this capacity is present in several approaches to CT, as well as in different theories of thought. What is missing from these approaches and what needs to be incorporated into a CT theory as rationality is consideration of the literature on heuristics and biases, particularly tests that can be used as tests of CT/rationality (Stanovich & Stanovich 2010;[59] West et al. 2008). The approach proposed by Stanovich thus makes it possible to consider ideas for measuring CT/rationality: to assess the degree of resistance to bias, independence of thought from these influences and the influence of prior beliefs, but also to assess resistance to other forms of bias described in the literature on heuristics and bias (Stanovich & Stanovich 2010[60], West, Topiak & Stanovich 2008). Developing CT - i.e., developing rationality - thus means developing and strengthening Type 2 thought processes, and evaluating CT thus means assessing the extent to which the subject uses Type 2 processes instead and is able to bring Type 1 processes under control. Indeed, according to Stanovich, the heuristics and bias approach, with all its empirical support, is theoretically dependent on a dual approach to thinking. In the latter, type 2 processes - cognitively costly, serial, consciously accessed, often verbalized - have the function of overcoming type 1 processes - fast, parallel, inexpensive, based on old and automatic associations, and capable of producing irrational responses because of the context in which they are activated.

A fundamental cognitive function for the exercise of rationality thus becomes that of[61] "inhibition", which is studied in the context of studies concerning executive functions. Inhibition is the mechanism that allows type 2 processes to bring type 1 processes under control and take their place (Stanovich & Stanovich 2010).

From his dual theory[62], Stanovich draws a further consideration regarding rationality/CT. Rationality and CT would belong to a different domain of intelligence as measured by current standardized tests. Indeed, measures of intelligence tend to be based on measures of efficiency in information processing in optimal situations: the subject knows that he must give the best possible answer and therefore use his best algorithms. The intelligence tests therefore implicitly put him/her in a situation where he/she uses his/her type 2 processes, and then measure his/her performance and efficiency in this implementation. On the other hand, a rationality test is a test conducted in a typical (vs. optimal) situation, which allows to evaluate if the subject implements type 2 processes while the situation allows, or even directs towards, the use of automatic and inexpensive processes. Rationality tests must measure this cognitive willingness to make cognitive effort, and the willingness to think rationally: for example, if the subject tends to seek information before forming an opinion, to calibrate the strength of his or her opinions on the basis of the available evidence, to think about consequences, to think before giving an answer, to weigh the pros and cons of a decision, to self-regulate with respect to its objectives. Conversely, intelligence tests are not about the ability to change one's mind in the face of evidence to the contrary, to direct actions in order to achieve a goal, or to get new knowledge (Stanovich & Stanovich 2010).

Stanovich thus reprinted Type 2 processes in two subsequent typologies: algorithmic thinking and reflective thinking[63]. The first is characterized by the ability to implement algorithmic-type processes, the second by the willingness to do so, i.e. to take a reflexive attitude when faced with a new problem. According to Stanovich, several studies show that the results of IQ-type cognitive tests and tests of resistance to bias, particularly belief bias (the bias of prior beliefs), are decorrelated. Thus, we cannot attribute success or failure in rationality tests solely to the measurement of intelligence (Stanovich & Stanovich 2010). For Stanovich, the exercise of rationality requires three different forms of ability:

- algorithmic capabilities must be present;
- a tendency to block and override automatic responses should be present when these are suboptimal;
- knowledge that can be retrieved from memory and that will help in problem solving (what Stanovich calls “mindware”).

IQ tests only measure the first form of these abilities [64]. Good results in this type of test do not exclude poor results in terms of rationality, for example in bias-based tests such as the "bat and ball test". Stanovich then speaks of "dysrationality": the inability to be rational despite a normal IQ. The reasons for this are defects in the other abilities necessary for rationality: the subject prefers less efficient but also less expensive solutions, or does not have the necessary mindware (Stanovich & Stanovich 2010).

Several aspects of Stanovich's approach are original to other CT theories. In particular, the notion of epistemic rationality is convincing for a narrow definition of CT. The idea of manipulating the environment to make up for motivational deficits is also original and opens the way to a broader vision of rationality - distributed over the subject and its environment. However, the dual theory of thinking and the program of heuristics and biases that underlie Stanovich's approach are subject to several criticisms: the rationality described by Stanovich would be an ideal rationality, disregarding the reasons that make normal cognitive functioning globally well adapted to solve the problems of everyday life. The difficulty to find the motivation to overcome the "errors" of system 1 and to substitute them - through reflection - the more algorithmic processes of type 2 could moreover be explained by this global functionality of rapidly available solutions. Stanovich does not consider that it is possible that systems and mechanisms for self-correction of errors are also present in an automatic, implicit way, and that it could be useful to develop them (without necessarily going through explicit, reflexive procedures). Stanovich emphasizes the importance of knowledge and strategies for the exercise of rationality/CT. However, he seems to view such knowledge and strategies as general, applicable in a variety of contexts, and not dependent on domain content. The difficulty of transferring a strategy from one content or context to another is therefore not considered by Stanovich.

Robert Sternberg: CT = intelligence

Sternberg attributes the birth of the modern movement of critical thinking to John Dewey and celebrates the richness of this movement, which draws on three different traditions: philosophical, psychological and educational. He considers philosophical approaches useful as guides but also as limited from a practical point of view. He appreciates the contributions of the educational approach - which he associates with figures such as Bloom, who have been able to propose theories of thinking and learning that take into account ecological observations, in the classroom, and not only (as is the case for psychologists) the results of somewhat artificial tests conducted in the laboratory. Finally, the usefulness of the psychological approach lies in its ability to shed light on the mental processes that humans use when they seek to think critically even if they are in suboptimal conditions: information is limited, time is limited, their memory is not perfect.

Sternberg henceforth proposes a taxonomy of CT abilities based on a psychological approach, in connection with his theory of intelligence. This taxonomy includes three types of capacities involved in CT: meta-components, performances, knowledge acquisition processes. Meta-components are higher-order processes used to plan, monitor what is done and evaluate the results of what has been done. For example, these components include the ability to identify the existence of a problem, to understand its nature, to plan in an orderly manner in order to solve it, to create a coherent strategy, and so on. (Sternberg 1986[65]). In addition to these higher-order components, there are various forms of reasoning and abilities that allow the instructions of the metacognitive components to be put into practice: spatial, causal, deductive reasoning, reading ability, etc. The performance components are lower-order processes. These abilities intervene in a variable way depending on the situation. An exhaustive list could not be produced. For example, there are specific components that enable reading, others that enable inductive reasoning, etc. (Sternberg 1986[66]). The third cognitive pillar of CT is represented by the abilities that enable the acquisition of new knowledge or learning of new procedures. The acquisition of knowledge and procedures is in turn based on three main components: selective encoding, combining information in a selective and relevant way, and comparing acquired knowledge with new information to be learned. These same components are the ones that allow for novelty, a fundamental condition for the exercise of CT (Sternberg 1986[67]). CT is therefore a set of mental processes, strategies, representations that we use to solve problems, make decisions, learn new concepts. Beyond the more general, metacognitive elements, the abilities required for CT vary according to the task, the situation and also the subjects.

From a practical educational point of view, Sternberg is aware of the importance of promoting the transfer of acquired abilities through a pedagogy that maximizes the chances of applying CT to real-life situations. He proposes to use concrete examples to illustrate the methods of critical thinking, varied exercises for training, including practical exercises on everyday issues but also on more academic issues. In a pedagogical manual, he proposes in a more general way to improve the intellectual capacities of the students, which he calls "intelligence". The manual focuses on the three components of the approach described, but also includes sections on emotional aspects and motivation to use one's intelligence. In terms of tests to measure CT, Sternberg has developed his own psychological test: *the Triarchic Test of Intellectual Skills*. This test does not separate, according to Sternberg's approach, CT and intelligence and seeks to measure meta components (planning, monitoring, evaluation), performative components (inferring relationships, applying relationships, comparing higher-order relationships between domains), knowledge acquisition components and the ability to cope with novelty (distinguishing relevant and irrelevant information, combining information logically, using acquired knowledge to learn or understand new concepts), automation of information development processes and adaptive flexibility.

Tim Van Gelder: CT as a form of expertise

Van Gelder, a philosopher by training, very interested in the cognitive sciences who consults for companies and as a researcher at the University of Melbourne, has published only one article on the topic of CT, but is still often cited in references concerning the psychological approach to CT. Van Gelder emphasizes both the value and the limitations of turning to cognitive science for guidance on how to teach CT: cognitive science provides a better understanding of how the mind works, but it is a fundamental area of research that does not provide practical guidance and empirical data on what works and what doesn't; and cognitive science has (as to 2005), paid little attention to higher-order thinking skills, including CT.
The lessons we can draw from cognitive science are, therefore of a fairly general nature: the need to practice CT capabilities, but also the need to acquire theoretical knowledge, the difficulty of acquiring expertise, and the need to practice, for transfer, the existence of confirmation bias. The only specific indication concerns the usefulness of practicing techniques such as visual representation of arguments, a practice that Van Gelder also encourages (Van Gelder 2005[68]). Van Gelder's fundamental theory is that CT is not natural. On the contrary, we could compare it to the practice of ballet: a highly artificial, expert activity that requires a long, dedicated, expensive exercise. Nothing in our evolution has favoured the kind of rational, logical reasoning that is characteristic of CT (Van Gelder 2005[69]). Van Gelder's position is certainly somewhat caricatured, because the CT, like other capabilities, could have a natural basis that would make possible the further, possibly cultural development of expert capabilities. The "just how much" it is useful to have CT capacities to survive in the environments of our evolution is not a question that we can answer in an

abstract way, and also depends on the capacities that we consider fundamental to CT. Van Gelder does not provide a definition of CT.

A further difficulty identified by Van Gelder in exercising CT is that it is a higher-order ability (he then compares CT to playing tennis), which is based on exercising other, more basic abilities. For example, in order to exercise CT in the interpretation of a letter, one must be able to read, etc. These difficulties explain why learning to think critically takes a lot of time, and Van Gelder compares this to the time it takes to master a second language. For all these reasons, there is no magic potion to become better critical thinkers. However, Van Gelder suggests several strategies. The first is practice. He draws on the literature on expertise, especially the work of Karl Anders Ericsson, and thus integrates into the field of CT Ericsson's advice to develop expertise in fields as different as chess, sport, and the practice of a musical instrument. Notably, the practice must be deliberate: assumed and voluntary. It must be designed for transfer: for example, it is a matter of practicing exercises rich in content in different contexts, but then encouraging abstraction to identify common features - even if there is no recipe on how to encourage transfer. The second is theory: for example, mastering vocabulary, the principles of argumentation, the laws of logic. Theory develops an understanding of what is done in practice. This essentially means dedicating courses to CT, which include theory and practice in equal parts. Van Gelder is convinced of the advantages of spatially representing arguments in order to better master their relations by visualizing them. He therefore suggests using this type of external tool to improve CT's argumentation capacities. Finally, Van Gelder draws attention to an additional difficulty compared to those mentioned in his first point: we are prey to illusions, biases, cognitive errors that occur silently and belong to our cognitive apparatus, for natural reasons but which can be culturally reinforced. The deepest and most pervasive bias is, according to Van Gelder, the one that leads to the preservation of one's beliefs and opinions, called "confirmation bias".

Daniel Willingham: CT depends on content

Daniel Willingham is a cognitive psychologist, expert in education. His analysis of CT is often cited as a pessimistic approach to the possibility of educating this ability. Willingham is indeed - within the framework of the panorama we have just outlined - the only critical thinking theorist who is strongly sceptical about the possibility of improving it, or at least of being able to improve it in a general way and through general education. The closest approach to this vision among those encountered so far is represented by the work of the philosophers Sharon Bailin and Mark Battersby (Bailin & Battersby 2018). Both philosophers, like Willingham, also strongly emphasise the dependence of thought processes on the content on which thought is engaged. Their approach to the education of critical thinking is therefore disciplinary and aims at improving thinking skills in one area, or at least in relation to certain contents. Bailin and Battersby criticize the desire to improve CT *through* formal logic courses in favour of the development of argumentative debates on specific themes. According to them, it is not a question of learning how to defend oneself (CT as intellectual self-defence) against poorly formulated arguments, but of learning how to look for the facts behind the information, how to choose one's sources, how to use epistemic norms adapted to the context and the contents under discussion, and how to adjust one's confidence on the basis of these evaluations. It is notably by comparing different positions, *through* debate and discussion - a method that Bailin and Battersby call "*inquiry*", "investigation" - that CT can be improved. The authors stress that investigation is contextual: the analysis of arguments depends on the content of the arguments, because, for example, the norms and standards for assessing the correctness of an argument may not be the same from one content to another. CT is therefore the ability to use epistemic norms appropriate to context and content to defend or attack an argument, justify oneself, or criticize a position. Such standards allow for the assessment of the quality of sources against standards, for judging causal arguments or reasoning involving statistics (Bailin & Battersby 2018).

Willingham adopts a more descriptive definition of CT, based on mobilized cognitive abilities. Similarly, his reasons for maintaining a skeptical stance towards a general teaching of CT derive from psychological, rather than epistemic, considerations.

For him, CT abilities belong to three types of mental activities that we use every day, but not necessarily critically: reasoning, problem solving, and decision-making. The adjective "critical" indicates that the processes of reasoning, decision-making, and problem solving are carried out voluntarily (autonomously), efficiently (unaffected by biases and other natural cognitive limitations), and creatively (not remembering a solution) (Willingham 2007[70]). So the problem is: can we teach people to think? The answer is: If we could, it wouldn't be like teaching or learning to ride a bicycle (Willingham 2007[71]). Willingham is pessimistic about the possibility of teaching CT because he considers that thinking is not a *skill*. On the contrary, he explains that thinking is strongly dependent on the content of knowledge. Therefore, learning general strategies does not guarantee to be able to apply them in concrete cases, on particular contents (Willingham 2007[72]). The dependence of content thinking is also revealed in the difficulty of detaching oneself from the superficial, concrete aspects of a problem (its factual content) to try to glimpse its "deep structure". When we seek to understand a new problem, we use the knowledge stored in our memory, but also the context. This makes understanding faster, but anchors it to the contextual content. It is these considerations that lead Daniel Willingham to assert that teaching - generally speaking - to think or think critically is impossible.

However, several programs are aimed at developing critical thinking or thinking in general. Willingham wonders about their results. Even when these methods appear to provide positive results, empirical evaluations have several methodological limitations: while we can say that many of these methods meet their "internal" objectives (learners learn to solve the types of problems they encounter in the program), it is more difficult to say whether the effects of these methods are generalized and transferred to "real life", or whether they are likely to be sustained over time. In some cases, it is also difficult to know whether the positive effect is due to the method itself or to other conditions. Willingham offers a brief analysis of the effectiveness of the better-known programmes for teaching critical thinking *per se*. He summarizes the common characteristics of the most widely used methods in three points:

- they presume the existence of skills that can be practised independently of context and content, CT teaching takes place outside a disciplinary framework;
- some are of long duration (three years, several hours of instruction per week);
- They all use examples of critical thinking and then ask to apply the strategies learned; some use abstract problems (*Ruven Feuerstein Instrumental Enrichment*), others use mystery stories (*Martin Covington Productive Thinking)* or group discussions on everyday problems (*Edward De Bono Cognitive Research Trust*: CORT).

Studies that measure the effects of these interventions have several limitations (Willingham 2007[73]):

- in some cases, students are evaluated only once, so we can't know if the effects last;
- in some cases, there is no control group, or the control group does not carry out an alternative activity (passive group) ;

- in some cases, there are no transfer measures to actual situations or to situations different from those used in the investigation ;
- only a small proportion of these studies have been subjected to the scientific process of publication and peer review;
- When we talk about the positive effects of these methods, we are not necessarily talking about the same thing as critical thinking, or a notion that specifically relates to critical thinking.

Despite all these difficulties and limitations, reviews of the literature or texts on critical thinking often explain that educational interventions for critical thinking have positive effects (Willingham 2007[74]). Is there no hope of transferring skills from one context to another - including 'general' skills such as the ability to solve a certain type of problem? Willingham points to two conditions for success: familiarity with deep content and knowing that one has to look at deep content. Familiarity depends on repeated, automated practice. The second strategy is metacognitive: knowing where and how to look for the right strategies in one's memory, which presupposes that it is at least necessary to want to and think about doing it. Finally, it is necessary to possess knowledge and to know how to mobilize it in practice. Without this, we know what we should do but we do not know how to do it[75]. The conclusion is that, although we have natural abilities, such as the ability to reason about causes and conditional probabilities (an ability that is, in intuitive form, limited from early childhood), we can continue to make mistakes in the use of conditional probabilities and causal reasoning even in the presence of more sophisticated abilities and knowledge. Thinking more critically depends very much on content, which is also true for more specialized forms of "expert" thinking, such as scientific thinking. Even in the case of scientific thinking, success depends not only on knowing procedures, strategies, but also on knowing when and how to apply them (Willingham 2007[76]). Critical thinking instruction, therefore, depends in part on teaching students strategies for "better thinking" and, in large part, on how and when to deploy these strategies.

Seemingly counter-intuitively, the development of CT capabilities is highly dependent on the acquisition of domain knowledge. Transferring competence on known content is easier for two reasons: first, because we know what aspect the competence in question takes on in the subject matter; second, because domain knowledge makes it easier to recognize the deeper structure of a problem and to move beyond the superficial examination stage (Willingham

2007[77]). Willingham cites cases of success in CT teaching where students have been exposed to strategies and repeatedly applied them. However, these are examples internal to a certain content area. Skepticism remains when it comes to imagining CT curricula aimed at general transfer, the application in all possible situations of such general strategies as analysis, synthesis, etc., as these strategies are simply not the same from one context to another. The objectives of CT teaching, to be realistic, must therefore be internal to a field. There is no substitute for domain expertise when it comes to thinking critically, recognizing the deep structure of a problem and thinking about appropriate strategies (Willingham 2019[78]). In the end, Willingham proposes a four-step plan for teaching CT (Willingham 2019[79]):

- identify what counts as critical thinking in each area: in history we do not consider facts in the same way as in science, for example ;
- teach the corresponding skills and strategies explicitly ;
- identify the important domain contents to be mastered in order to be able to think in the domain in question ;
- plan for the long term and re-expose learners to the skills to be learned, have them practice (for three to five years).

Common features of psychological and philosophical approaches

A relatively clear division is drawn in the literature between philosophical and psychological approaches: the former are rather normative and the latter rather descriptive (Lai 2011, Sternberg 1986).

While philosophical approaches indicate norms and standards to be respected in order to think well, psychological approaches rather provide lists of cognitive abilities to be developed or exercised in order to improve decision-making, make thinking more effective, or simply exercise higher-order forms of thinking that are opposed to other simpler forms that are less adapted to complex contexts. In other words, for those psychologists who have taken an interest in CT, it is a set of abilities used to achieve a certain goal or function, such as solving problems, making decisions, and learning new concepts. These abilities are considered necessary to provide appropriate or effective responses in complex situations where automatic, unthinking, formulaic responses cannot produce the desired result.

In reality, this distinction is less clear-cut: philosophical approaches include lists of abilities, just like the psychological approaches proposed by authors such as Diane Halpern or Robert Sternberg. And psychological approaches also include norms, as do philosophical approaches: Richard Nisbett talks about the importance of the rules of logic, the rules of statistical mathematics, strategies for identifying good evidence; Robert Stanovich contrasts rapid thinking processes with processes that can use learned rules; Diane Halpern points out the right way to use natural abilities to achieve more effective thinking, and she refers to criteria and norms. Thus, even if psychologists have not developed a reflection on the notion of criteria, the concepts of standards, norms to think well, strategies to help thinking are present more or less implicitly in psychological approaches.

The debate around the role of domain knowledge also cuts across the division between philosophical and psychological approaches. For example, Diane Halpern and Tim Van Gelder (e.g. Halpern 2007, Van Gelder 2005) consider CT to consist of general capacities. For Daniel Willingham, it is a set of capacities that are exercised on specific contents, which makes the transfer of knowledge complicated (Willingham 2007, 2019).

The debate about the role of knowledge also depends on the debate about the transferability of the CT (see Lai 2011[80]). The most optimistic authors are those who consider the CT as a general capacity (we have just quoted Halpern 2007, Van Gelder 2005 and we have cited Lipamn 1987 as an example for philosophical approaches). Others tend rather to see CT as content-related and thus more difficult to transfer (this is the case of Willingham 2007, 2019 but also Bailin 2002 among philosophers). Finally, some consider that the specificity of the CT is not an objection to its transferability (for example, McPeck 1990). From this debate arise the divergences between the different educational and testing methods of CT. This reason for divergence is in addition to the number and type of abilities considered fundamental in defining the CT, which has an impact on the abilities taught and assessed (Lai 2011[81]).

The following components are present in both philosophical and psychological approaches to CT:

- the common goal is to arrive at more satisfactory judgments or decisions (but the nature of this satisfaction varies from one author to another).
- both sides largely agree on the capacities or general skills to be cultivated in developing or practising CT, including analysis, evaluation and judgement of

arguments, assertions, evidence; the use of inferences and forms of inductive and deductive reasoning; reference to decision-making and problem solving (Lai 2011[82]).

- In particular, most authors associate the critical component of critical thinking with the ability or activity of evaluating information, arguments, assertions, and so on.
- Also, the concept of metacognition is often - not constantly - present. It can be interpreted differently from one author to another. It is always about explicit metacognition, self-reflexivity and monitoring of performance or the reasons used to justify it (Lai 2011[83]).
- Another common point between philosophical and psychological approaches concerns pessimism about the ability to naturally think critically (see Lai 2011[84]). The very idea of natural CT is never mentioned: CT is poor, even in the case of educated adult subjects. Hence the need for forms of education that are designed to develop CT (at the disciplinary or general level), not limited to the transmission of knowledge. Another consequence of this initial consideration is that the various authors - including psychologists - have not lingered on researching and describing the natural bases of CT that made his education possible, but they focused either on the goals to be achieved (philosophers) or on the obstacles to be overcome (psychologists). The only exception is Deanna Kuhn. She focuses her approach more particularly on one ability: metacognition. Moreover, Kuhn's approach is developmental, which means that she traces the different stages of the natural development of metacognition in order to give educational indications to go beyond the natural limits of this capacity and develop an expert form of it (see Lai 2011[85]). However, even in Kuhn's case, we can extrapolate that the true CT is only that which is considered expert and that natural metacognition is an imperfect basis for it.

2.2 Reasons for dissatisfaction with existing CT approaches

2.2.1 Definitions too broad to be operational

We indicated at the beginning of this section that the first reason for dissatisfaction in the existing CT literature is the lack of a consensus definition. The apparent agreement among different authors is largely due to the fact that definitions are often quite generic and that CT taxonomies include a wide variety of capabilities and provisions. For example, CT is often associated with skills and attitudes as varied as (see Ennis 2016 :

- Analytically and carefully consider any belief or opinion and look for reasons to believe; know the methods of logical and rational analysis of arguments and know how to apply them; adopt explicit criteria in the operations of thinking and use them to guide judgment in a variety of contexts; know how to conceptualize ; be able to analyze, synthesize and evaluate concepts and information; be able to use these skills to guide action; use reason to make better decisions, reasoning; be able to solve problems; possess metacognitive, reflective skills; possess algorithmic and logical reasoning skills; demonstrate clarity of expression, argumentative skills ;
- to engage in the voluntary exercise of all the abilities mentioned; to have a disposition to reflexivity, to the exercise of rationality; to act on the sole basis of reasons; to show a form of open-mindedness, a desire to be well informed, flexibility ; show a certain curiosity, an ability to recognize a lack of information, an ability to change one's mind and to suspend one's judgment; be always ready to engage in discussion; show a propensity to seek reason, to question and to question oneself;
- be sensitive to ethics and epistemic values, fairness; respect truth and respect the person you are dealing with.

Although convincing in the first instance, approaches using such taxonomies are unsatisfactory, particularly from an operational perspective. Indeed, if the CT concept were to cover all of the capabilities and provisions listed above, it would be too **broad**. Associating the CT concept with too broad a set of capabilities is problematic for assessment and education: it simply becomes impossible to address all of the dimensions related to the CT concept. Take the definition provided by the Delphi panel (Facione 1990) as an example. It states that CT is an ideal of making judgements, voluntarily, reflectively, and being able to justify those judgements on the basis of methodological, conceptual, or evidence-based considerations. In order to achieve this, it is therefore necessary to master all the abilities (and sub-abilities) of the CT, i.e. to interpret, analyse, evaluate, make correct inferences, explain, and self-regulate. These capabilities then include sub-categories: categorizing, decoding meaning, clarifying content, examining ideas, identifying arguments, analyzing them, evaluating assertions and arguments, searching for evidence, imagining alternatives, reaching conclusions, presenting results, justifying procedures, presenting arguments, self-regulation, and self-correction. Could a CT education program address all of these capacities? And, if so, what would be specific about this program in relation to CT, and different from an educational program that would aim to develop the full range of thinking, reasoning, problem-solving, oral and written expression, as well as metacognitive abilities of its candidates?

If it relies on too wide a variety of functions, and if the definition is too general, the concept of CT loses its **treatability**. Choices must therefore be made to reduce the range of skills to be taught to improve CT.

2.2.2 Definitions that are too vague and may lead to misunderstandings

The second reason for dissatisfaction relates to the **vague** nature of some of the capacities and provisions associated with CT, such as the ability to "change one's mind", "pay attention to the facts" or "remain open to the ideas of others". Again, we can cite the Delphi Report, which describes the critical thinker as curious, well-informed, trusting in reason, open, flexible, balanced in his or her assessments, honest in dealing with his or her own biases, cautious in making judgements, willing to take a step back, clear, orderly, diligent in seeking information, reasonable in selecting the best course of action, and willing to make judgements.e ction of criteria, focused on investigation, etc. (Facione 1990).

Without more precision, these indications lend themselves to be memorized as abstract concepts but they are not easily declarable in an operational way in very precise situations. Paying attention to facts or determining the value of evidence, for example, are difficult considerations to put into practice if we do not know what evidence, or even good evidence, is in the given context (see in this respect Willingham's 2007 critique of principled educational approaches to CT). If it is based on vague notions, the concept of CT thus loses its **relevance**.

A definition that is too vague can also have opposite, or even undesirable, consequences. The fact that CT is associated with vigilance against misinformation and a call for greater openness illustrates such paradoxes. Provisions such as the readiness to backtrack, if left unspecified, can be misinterpreted. Systematically changing one's mind is no more desirable than never changing one's mind. Indeed, reasonably speaking, we are not as ready to abandon our ideas when they are supported by numerous and solid proofs as when they are, on the contrary, based on impressions or vague intuitions. It is important to point out that the critical thinker is not always looking for information, but that he knows how to stop at a position in good confidence, if it is supported by good reasons and evidence. Otherwise, the critical thinker would spend his or her life constantly questioning his or her own or others' positions. The vague and abstract nature of these capacities and dispositions means that the notion of CT can easily slide into undesirable forms of relativism (doubting everything), paralysis of action (suspending judgment, questioning everything) and practical contradictions (remaining open, being suspicious). Existing definitions sometimes make it difficult to draw a clear line between CT and the attitudes of conspiracy theories that are characterized by a relativistic attitude and the construction of endless "argumentative mille-feuille" (Bronner 2013). If it is based on vague notions, then the concept of CT loses its **validity**.

2.2.3 Definitions that make reference to ambiguous concepts

Some definitions and taxonomies bring CT closer to concepts such as reason, reasoning, rationality, or reasonableness. This is the case, for example, of Willingham (2007), who identifies CT with three categories of mental activities: reasoning, making judgments or decisions, and problem solving. These concepts are no better specified and defined than the concept of CT itself. One needs only think of the debates over the definition of the concept of reason and rationality (see Mercier & Sperber 2011 and Mercier & Sperber 2017[86] for a discussion on this subject). In the developmental psychology literature, the word "reasoning" is associated with specific forms of representation and inference that concern objects, agents, quantities, geometry, and others (Spelke & Kinzler 2007). We also refer to causal reasoning (Gopnik & Schulz 2007), reasoning by analogy (Gentner 1989), and we look for its early origins in the cognitive development of the child. There would thus not be a single type of reasoning, independent of its content or means.

Most often, however, the term CT is associated with a specific meaning of reasoning, developed within the framework of the psychology of reasoning and decision-making, notably the "dual" model known to the general public as "System 1/System 2" and the associated "heuristics and bias" program model (this is the case of Stanovich & Stanovich 2010 and Nisbett 2015, for example). The idea behind the "heuristics and bias" program - for example, as expressed in the founding text of Kahneman and Tversky (1974) - is the following: under conditions of judgment under uncertainty, we use shortcuts that are limited in number, quick to implement and simple. These heuristics are generally useful but can sometimes be misleading. When a heuristic gives rise to a systematic error, we speak of bias (Kahneman et al. 1982, Gilovich et al 2002, Kahneman & Tversky 1996). Thus, CT is associated with the operation of a "System 2" that is rational, slow, algorithmic, often correct, and with the idea of inhibition or control of cognitive biases (debiasing*)*. This view has given rise to debate and criticism, particularly in the context of an evolutionary approach that emphasizes the adaptive nature of cognitive processes - by analogy with other traits selected over the course of evolution (see Cosmides & Tooby 1994, Gigerenzer 1991, 1996, 2004, Gigerenzer et al. 2008, Haselton et al. 2009, but this point will be developed later in the Report). Gigerenzer, Tood & The ABC Research group (1999), for example, contrasted the "heuristics and bias" program with "quick and dirty*"* solutions in which certain heuristics provide fast, economical and correct ("fast and frugal*")* answers. This framework views heuristics as part of an adaptive mental "toolbox", selected throughout our evolutionary history in response to specific tasks and problems that they efficiently enable. It would therefore be wrong to overemphasize the negative or biased nature of such solutions. The "heuristics and bias" program is also criticized because of the artificial nature of laboratory test situations in which biases are highlighted. If these situations are modified, the biases would simply disappear, or at least be weakened. For example, studies on cognitive biases highlight the limitations of probabilistic reasoning. However reformulating the problems posed in terms of frequency makes it possible to increa se the number of correct responses; this is because, in natural conditions, we spontaneously rely on the observation that an event is more or less frequent (Gigerenzer et al. 2008). Finally, adaptive evolutionary approaches criticize the "heuristic and biased" view for using only external norms to judge whether behaviour is rational or irrational. In the heuristic-bias approach, the norm is set based on the desirability of a certain type of behaviour. Whereas in the adaptive approach, we seek to establish whether the norm really corresponds to the objectives that the cognitive system responds to in order to increase the individual's selective value, under given conditions. Thus, while truth-seeking is a culturally important norm, there is no assurance that the solutions selected in evolution meet only such norms.

This selection was made under multiple constraints - including, but not limited to the search for truth - and in complex environments, which we often neglect in a classical "heuristic and bias" approach. On the basis of these considerations, some trends that are considered to lead to irrational decision-making - risk aversion, for example - can be reinterpreted as highly effective strategies depending on whether we consider one optimality rather than another (for example, minimizing the variance of earnings - except under conditions where strong earnings are needed - rather than simply maximizing earnings) (Cosmides 1989, Cosmides & Tooby 1996, Gigerenzer 1991).

The evolutionary approach can thus reveal the adaptive nature of *a priori* biased decisions, and the evolutionary pressures that have shaped the decision-making processes. Evolutionary criticism of heuristic and biased approaches also makes it possible to highlight the need to question the origin of errors and biases that affect decision-making and the formation of opinions and beliefs. This is in order to target realistic improvement strategies rather than simply admitting these biases and seeking to counter them. However, it seems excessive to oppose the two approaches "quick and dirty" and "fast and frugal" in too clear a manner. The two approaches recognize that natural cognitive functioning enables us to solve many situations effectively, but that it also has limitations and can even mislead us. The idea of bias and strategies to overcome or avoid certain biases is therefore not to be rejected in a CT approach, but it would be excessive - considering the existing criticisms - to reduce CT education to a fight against bias and CT itself to a way of thinking that would allow us to "debiate" (a vision present in authors such as Keith Stanovich, see Stanovich & Stanovich 2010).

2.2.4 Definitions that are too unrealistic

We find that existing approaches to CT often show a certain **lack of psychological realism**. Philosophical literature in particular often refers to a kind of ideal thinker without necessarily questioning the achievability of the objectives set for this ideal thinker, nor his or her "starting" capacities (Lai, 2011). As an example, we can cite the definition that emerged from the consensus-building work on CT (Delphi method) coordinated by Peter Facione (Facione 1990[87]).

The literature in psychology, for its part, does not bother to define the natural bases of critical thinking in developmental terms (with the exception of the work of Deanna Kuhn). The descriptive effort is often focused on identifying the obstacles (biases) that stand in the way of the "correct" response or the choice considered a *priori* as optimal, and on the subject matter.ition of remediation (this is the case of Richard Nisbett and Keith Stanovich in particular, see Nisbett 2015 and Stanovich & Stanovich 2010). However, an analysis of CT in terms of cognitive abilities made it possible first of all, of their specific en link with the CT, their limitations. None of the authors reviewed, apart from Kuhn (see Kuhn 1999), raises the issue of CT outside the educational framework. For the philosophers as for the psychologists we met, CT is therefore an objective to be achieved, a state which is born of artificial or at least learned strategies. We have seen that Kuhn, on the contrary, identifies a group of metacognitive capacities as the natural basis of CT, and then seeks to describe the ontogenetic development of these capacities, before any form of education, until they reach a plateau that requires an educational effort to be surpassed. His approach to the metacognitive development of the child is, however, dated. The problem is therefore to place CT in the current framework of knowledge on cognition.

A further aspect of the lack of realism in current approaches is the absence of references to certain "ecological" aspects of the exercise of critical thinking. Social aspects of cognition are thus poorly represented (exceptions are represented by Bailin & Battersby 2016; Lease 2011; Kuhn 1999). However, the exercise of critical thinking takes place in an often social context. We exchange information with others, receive information and pass it on to others. The information that comes to us from others (second-hand information) comes to us with reputational clues, considerations concerning the prestige of the source, its presumed expertise, its familiarity with us - hence its reliability (Origgi 2015). The arguments that we are led to judge exist in a contexte cultural and are the subject of ideological positions, linked to opinion groups (Kahan 2015). Removing CT from this context risks making it a concept **disconnected from its own reality**.

We also exercise our abilities in relation to others. We discuss our positions, we argue. We do not limit ourselves to carrying out our analyses, evaluations, reasoning in a situation of isolation, but often in a situation of exchange (see Mercier and Sperber 2017, Mercier and Sperber 2011, Trouche et al. 2016 for an argumentative approach to reasoning, in the sense that reasoning is a process that makes sense in the exchange of arguments).

The social dimension of CT becomes all the more important as this concept is often evoked in relation to phenomena of a cultural and social nature such as conspiracies, the circulation of false information and its possible influence on decision-making (fake news), phenomena of mistrust of knowledge based on solid evidence (the safe nature of vaccination, facts in support of global warming) (see Halpern 2013). In this framework, the problem of trust and who to trust (Oreskes 2019, Bronner 2013) arises, i.e. the ability to evaluate information sources, their degree of reliability, their skills, the possible levers used to convince, but also the factors that make one piece of information more attractive and easy to remember than another. Where do these capabilities come from? Are we naturally equipped with them? How do we trust them and why?

It is to answer these types of questions that we consider it necessary to base the theory of CT on a thorough analysis of natural cognitive functioning, to take into account developmental aspects, the ecological conditions in which the capacities of CT are expressed, their neurocognitive bases.

2.3 What is CT?

On the basis of considerations on the limitations of existing definitions and approaches, we propose a minimalist definition of CT. This allows to isolate the cognitive abilities directly involved in CT from those incidental to it. It will also be easier to estimate where an action to develop CT begins and ends and to identify assessment tools that will specifically measure CT.

As a first approach, we define CT as *the set of capabilities and criteria for assessing the epistemic quality of available information and for consistently calibrating our confidence in that information, with a view to making a decision, forming an opinion, accepting or rejecting a claim appropriately.*

Assessing the epistemic quality of information means asking whether the information has a good chance of corresponding to reality, therefore whether it deserves our trust: is it *plausible* the light of existing knowledge*?* Is it *relevant* in terms of good arguments? Is it *solid* in terms of evidence*?* Is the source of the information *reliable* – e.g., lacking private self-serving interests and competent in the domain?

Although this definition is narrower than most of those encountered, it is not at odds with them. Indeed, the notion of evaluation is present in most of the approaches encountered. Suffice it to cite the Delphi report, which inclsudes evaluation among its six components (Facione 1990), and the psychological approaches of Halpern and Kuhn, which explicitly refer to evaluation as the critical component of critical thinking (Halpern 2013) and the evaluation of supporting evidence as a criterion for distinguishing between opinions and knowledge (Kuhn 1999). A definition of CT based on the evaluation function of the quality of the epistemic nature of the information therefore has the triple advantage of being **specific** and **minimalist** and, and at the same time, to be **a continuation of existing approaches**. As a result, it is not simply *another* CT definition or a definition of *something other* than CT.

However, some distinctions are necessary, particularly in order to identify CT assessment tools:

- It is important to consider the mechanisms for assessing information and those for making decisions separately. When a subject makes a decision, he or she does so under the influence of a very large number of constraints, and the ability to evaluate information fairly is only one aspect of the decision-making process. The benefits of integration into a social group, for example, may take precedence over the acceptance of reliable knowledge (see, in this regard, the literature on motivated reasoning and the socio-cultural reasons behind choices such as vaccination, or the acceptance of the theory of evolution and climate change: Kahan et al. 2011, Kahan 2015, Kahan & Stanovich 2016, Kahan et al 2010). It is therefore desirable to separate these two aspects, although in practice this may not be easy to do (the existing literature does not necessarily make this distinction. For example, (Halpern 2013) considers CT to coincide with capabilities that allow for more effective and efficient decision-making or problem solving);
- when we are faced with an opinion or decision, we can rely on an assessment of the evidence that support it. The information available is reviewed to a greater or lesser extent. We can then justify our choice on the basis of arguments drawn from this assessment, link the arguments together coherently, or even present them eloquently to an interlocutor. The process of presenting and explaining our choices on the basis of arguments is, however, a separate process from the initial assessment of the quality of the information. It does not, therefore, fit into our narrow definition of CT, but can be seen as an ancillary capability. Recall that, in the case of the definition provided by the Delphi panel, CT includes both evaluative and explanatory capabilities. Similarly, the Delphi panel included interpretation, analysis and inference in its definition (Facione 1990). In this case, these are abilities associated with the evaluation of the quality of information, but which can also be part of other thought and reasoning

processes, and can be separated from the evaluation itself. Inference capacities, for example, are omnipresent in cognitive functioning (see Mercier & Sperber 2017). Interpretation skills are often prerequisites (for example, the interpretation of a text is a prerequisite to the evaluation of the epistemic quality of its content, but the two processes are distinct). Those of analysis of the logical construction of an argument are complementary. For this reason, unlike the Delphi panel, we do not include these capacities in our specific definition of CT.

- Several authors refer to the notion of metacognition or self-regulation (e.g. Kuhn 1999 Halpern 2013, Facione 1990, Ennis 2016). The use of this notion is not unequivocal in the CT literature and elsewhere. Kuhn, for example, uses it to speak of meta-knowledge, the use of strategies, and also as a form of naïve epistemology (Kuhn 1999). In addition, the understanding of metacognitive abilities and their development has evolved recently. Only a thorough review of the recent cognitive science literature, therefore, could tell us the extent to which CT as an assessment of how much confidence we can place in a certain piece of information depends on some form of metacognition.

On the basis of the definition provided, it is now possible to turn to the cognitive science literature to identify the natural bases of CT.

3. The cognitive bases of CT

3.1 Naturalizing CT

With few exceptions, philosophical and psychological approaches tend to ignore the "natural" and developmental dimension of CT. CT is defined as a goal to be achieved, in particular through dedicated education, and current approaches therefore focus more on what hinders it and how to develop it than on its possible natural basis. The implicit message that emerges is that CT is not "natural", but a conquest of education.

We often hear our fellow citizens (and ourselves) described as being gullible beings, prone to to fall into the traps of *fake news* and the meshes of persuasion, without any natural endowment to protect themselves from it. Of course, we can all offer anecdotes to support this impression, and the media attention paid to the *fake news* phenomenon only feeds this observation of fragility in the face of manipulation or dubious beliefs. But are we really - by nature - completely devoid of CT?

We will search the relevant literature for answers to the following questions: Are there natural cognitive functions/processes that enable us to evaluate information and to calibrate our confidence in the information thus evaluated, with a view to making decisions? What is the developmental path of these functions/processes? What are their limitations? More generally, the question guiding this second part is: Are there natural bases for CT? If so, how do they develop naturally?

The narrow definition of CT that we proposed at the end of the first part makes it easier for us to start looking for the cognitive foundations that make CT possible.

All these questions might seem purely theoretical. Would we limit ourselves to trying to learn more about CT as a product of human cognition? In reality, the objective of the naturalistic approach presented here is operational - as with other philosophical and psychological approaches to CT. Identifying the natural underpinnings of CT allows us to provide an objective starting point for CT education. Identifying the natural bases of CT also allows us to establish more objectively the limits thereofand thus to design pedagogical actions or facilities that will allow the overtake.

Natural CT, which corresponds to these bases is indeed limited and does not always serve us well in all circumstances. Accusations of credulity are not entirely false, by the way, and eastward force to note that we are not naturally equipped to face contemporary challenges, the multiple forms of information flow, to distinguish the most solidly founded knowledge from popular opinion widely disseminated.

The same applies to our ability to build objective knowledge. The literature in developmental psychology is rich in examples illustrating the presence of a natural form of scientific reasoning in children. The baby is already reasoning about causal relationships, has conceptions and expectations about the laws of physics, distinguishes the world from the living, animate world from the inanimate world, and assigns mental states to others. He explores knowingly, guided by a curiosity that is not random but rational, at least from certain points of view. He possesses a tooling allowing him to update the hypotheses he makes day after day thanks to his experience and prior knowledge. However, even though he is so well endowed from birth, this baby is not a scientist. Understanding the explanations of professional science, abandoning the more intuitive explanations that come from his baggage of intuitions and experiences, is far from simple. Even in adulthood, human beings struggle to obtain an advanced understanding of reality. Experts in the field of building knowledge about the natural world - scientists by profession - are therefore endowed with a wealth of experience and expertise – a cognitive toolbox - in order to verify our intuitions and to control some of our biases that hinder an objective knowledge of reality. Scientists can thus arrive at objective, though often counter-intuitive, explanations. The methods are the result of a long history and of our cultural evolution. In the course of its history, human beings have indeed produced artifacts, strategies, to equip the natural scientific mind, already present in the baby - our natural endowment, forged by millions of years of evolution. The result is an expert form of knowledge of the natural and social world.

We can design the development of CT following the same model as we have just adopted for the scientific mind. The baby comes into the world with natural CT abilities, which he or she exercises in daily life to cope with information and decide on hypotheses for decision-making. The tasks that the child has to perform are not drastically different from those faced by our ancestors: the stresses of those days have shaped our cognition, making the baby capable of solving those tasks, then as now.

However, there is no reason to believe that our CT is flawless: on the one hand, natural selection does not optimize all traits simultaneously, nor does it optimize certain traits beyond a threshold value (because this optimization is not always possible no longer translates into adaptive value); on the other hand, our world today is different from the one for which evolution has prepared us. We need only think of the systems for the circulation of information to realize this.

Our demands have also changed: faced with a constantly changing social and cultural universe, we are probably looking for more objectivity and truth than in the past, which would be unnatural if we limited ourselves to considerations of the immediate adaptive type. We are natural and cultural beings, subject to this dual evolution.

This justifies that we seek to artificially develop our natural EC capabilities. **Developing natural CT then means equipping it with strategies and criteria to deal with situations different from those for which our evolutionary history has equipped us.** Natural CT thus becomes **advanced CT**. Professionals struggling with the evaluation of specialized information - in medicine, history, physics... - need equally specific tools. This translates into different forms of **expert** CT, capable of meeting the challenges posed by the need to acquires Increasingly sophisticated and specialized knowledge.

From this framework, we review the literature in cognitive science (with a focus on developmental psychology, and cognitive neuroscience) in search of the natural cognitive *building blocks* of natural CT. We will end with an excursion into the psychology of reasoning and in the psychology of evolutionary in order to trace the limits of natural CT. Finally, we will provide practical, operational guidance on how - from prior considerations - which we could use, in an educational context, to develop CT into more advanced forms.

3.2 The cognitive bases of CT: epistemic vigilance

Let's imagine we have to decide where to eat in a new city. Instead of trying out all the restaurants, we will probably try to get second-hand information, for example, by asking someone which is the best restaurant in the area.

This tendency to acquire information through others is universal and finds its *raison d'être* in saving time and energy as well as in the reduced risk-taking it represents, compared to taking information directly. We could give many examples: following the advice of someone who tells us not to eat a certain type of mushroom, for example, avoids the risk of poisoning ourselves. However, in order for the information provided to be useful, we must make sure that our informant is not lying to us voluntarily and that he or she knows what is going on. Indeed, dependence on others to acquire vital information carries the major risk of being misinformed[88]. The stakes are so high (being able to benefit from second-hand information without taking excessive risks) that it seems plausible, from an evolutionary point of view, that the selective pressure exerted on our cognitive functions has given rise to mechanisms capable of filtering information and protecting us from misinformation - even imperfectly. This is done in a cost-efficient and at least approximately correct way (Harris & Corriveau 2011, Sperber et al. 2010). All this could only be an evolutionary speculation - an attractive but unverifiable "just so story" - if it wasn't for the fact that certain protective mechanisms seem to be at work in our cognitive functioning, and this from childhood.

Indeed, empirical studies show that we have a number of mechanisms that have been studied, particularly in children, from a developmental perspective (Sperber et al. 2010, Harris, 2012) and grouped under the term "epistemic vigilance" (Sperber et al. 2010[89]). These mechanisms help us judge both the relevance and the epistemic quality of second-hand information - and thus to conclude whether information provided by others is likely to be accurate (or whether our informant is likely to be a good informant). In this way, we can quickly determine whether information provided by others is likely to be accurate or not. These assessment processes are not necessarily conscious, thoughtful or costly, and they can rely on simple clues. The result of their action is a more or less strong state of confidence in the information given, which translates from a behavioural point of view into a greater or lesser availability to use it to form an opinion/belief or make another type of decision. These processes have been studied, particularly in children, from a developmental perspective (Harris, 2012).

3.2.1 Mechanisms of epistemic vigilance

Epistemic vigilance applied to second-hand information is exercised at two levels: it concerns both the content - what to believe - and the sources of the information - who to believe.

The first group of mechanisms concerns the content of the information transmitted, regardless of its source. They consist, for example, in assessing the coherence with our prior knowledge (or background beliefs) or with other findings that we can make directly (and in which we trust). They are therefore mechanisms for assessing the plausibility of the content (*plausibility* check) and its relevance[90].

The second group of mechanisms applies to the source of the information (*trustworthiness*). This can be a source of misinformation for two distinct reasons: it is in possession of good information but wants to deceive us (*malevolence*) or it is benevolent but ignorant (*ignorance*). So there are mechanisms to deal with both possibilities[91].

An example helps to understand the articulation of these mechanisms. If we ask a passer-by for the way to the polling station in our neighbourhood, our attention will be drawn to the fact that he or she is wearing a distinctive sign of the League of Abstentionists or that he or she seems to be improvising (evaluation of the source: malicious or ignorant) or that he or she mentions a path that includes detours of several kilometres (evaluation of the content: not consistent with what we know about the network of polling stations or the geography of the neighbourhood).

In the literature, the idea of epistemic vigilance is based, in addition to evolutionary considerations, on observations made in particular on young children. Studies show that young children show *selective trust*: they use criteria very early on that allow them to choose an informant among others in a non-random way. They generally show a preference for caring and competent adults[92].

3.2.1.1 Mechanisms of selective trust in children (source-directed vigilance)

A significant number of studies show that from the age of three onwards, children use (implicit) criteria in order to choose one informant among others in a non-random way (*selective trust*, Harris, 2012): on the one hand, criteria that allow them to select an informant after they have been able to interact with him/her (familiarity of the informant, clues that make it possible to infer the informant's past competence); on the other hand, criteria that allow to selectively pick an informant even when they are in contact with him/her for the first time (generic competences observed in a direct way or through prestige or consensual cues, indices of goodwill, such as belonging to the same language group) (Harris & Corriveau 2011).

In a typical experiment, the young children watch two videos. In each video, an adult refers to objects or their functions. The two adults provide different names for one of these new objects. For example, Adult A calls the object a "*snegg*" and Adult B calls the object a "*hoog*". The child is then asked: What is the name of the "unknown" object? Adult A is an adult who is familiar to the child (a teacher at her school) and adult B is a teacher from another school, unknown to the child. Result: Children aged three to five show a preference for the familiar adult, A if they belong to school A, B if they belong to school B. Conclusion: the children did not choose their informant at random but used an index, that of familiarity, to make their choice. But familiarity is not everything.

In another experiment, conducted with three- to four-year-old children, the two informants (unfamiliar to the children) were first asked to name common objects known to the children. One of the two adults makes a mistake in naming these objects (for example, it calls "ball" a cup). If the children then have to ask for the name of a new and unknown object, they will decide to turn to the adult who has not been mistaken about the known objects; if both adults provide different names for the same unknown object, the children will choose the name provided by the most knowledgeable adult. They therefore orient their choices on the basis of the skills shown by the adult, and follow the most competent adult.

A third experiment compares the effects of familiarity and accuracy. In this case, one of the two informants is familiar, the other is not; the one who is mistaken about the common objects is, for one group of children, the familiar informant and for the other group the unfamiliar informant. For three-year-old children, the fact that the familiar adult makes mistakes is not enough to take away their trust. But for five-year-olds, being correct becomes a more important condition of trust than familiarity. Not only are children sensitive to the adult's specific competence, but they are also more sensitive to the adult's specific competence. Not only they prefer adults who can correctly name common objects, but they also tend to prefer those who show generic skills, such as using correct grammar (in the presence of an adult who pronounces "*a shoes*" and one who correctly says "*some shoes*", the children prefer the second as an informant).

A fourth group of experiments shows that young children are sensitive to social cues when choosing their informant. For example, they seem to prefer an adult with a native accent to an adult with a foreign accent[93].

In a fifth group of experiments, children are experimentally confronted with a case of disagreement between two adults, one receiving approval from other adults, the other not. The preferred adult is the one who receives the approval of the others. Through this battery of studies, we can show that children are selectively attracted to adults who are "culturally conformist", because they receive the approval of the majority[94]. Morgan, Laland & Harris (2015) tested the influence of the degree of consensus: they presented children aged between three and seven with images with varying numbers of dots and asked them to estimate which of two images had the most dots. The task varied in difficulty because of the ratio of dots between the two images. Three-year-olds were influenced by social consensus only if it was total; otherwise, the influence of the social context was random. From the age of seven, children also responded to non-absolute majorities and proportionally to the level of consensus.

In conclusion, children have two main categories of heuristics that allow them to gauge the reliability of their sources of information: on the one hand, heuristics that are used to select an informant after they have been able to interact with him/her (familiarity, past competence); on the other hand, heuristics that allow them to choose even when they are first confronted with the informant (generic skills, linguistic affiliation, consensuality[95]).
The criteria used by children change as they develop. Certain evolutionary trajectories have been highlighted, such as the increasingly important influence of the criterion of expertise versus familiarity, or sensitivity to more mixed forms of consensus and the ability to adapt one's confidence according to the degree of perceived consensus. However, the study of the development of selective confidence mechanisms and criteria during childhood is still work in progress.

3.2.1.2 Naïve first order epistemology (vigilance directed at content)

A broad tradition attributes to children the ability to distinguish between beliefs (representations that may be false) and reality only from the age of four to five (Wellmann, Cross, Watson 2001). Recently, this dogma has been shaken by studies showing that this ability is mastered implicitly, non-propositively, already around one and a half to two years (Scott & Baillargeon 2017). A proposed solution to this apparent contradiction is to distinguish between implicit and explicit forms of attributing beliefs and other mental states to others (Rakoczy 2017; Apperly & Butterfill 2009). In this perspective, Mascaro and Morin (2015) studied children's ability to deal with the notion of falsity before the age of five. Through their analysis of the literature and several experiments, they showed that it develops gradually and that two-year-olds already have the ability to consider an assertion as false. However, before the age of four to five, they have trouble memorizing false assertions, and they tend to correct them unintentionally. A possible interpretation for this type of inconsistency is that, although they are able to understand the distinction between true and false, younger children are more confident in their informants and tend to attribute true beliefs to them, and therefore tend to implicitly correct them if these informants manifest clearly false beliefs. Children therefore demonstrate a form of early naïve epistemology that allows them to evaluate information content as true or false[96].

The ability to appreciate the value of evidence in support of an assertion has been studied more in research on causal and scientific reasoning in children and adolescents (Zimmerman 2000; Morris et al. 2012) than in studies of epistemic vigilance. E.g., Gopnik and colleagues (2001), Schulz and Gopnik (2004) and Schulz, Gopnik and Glymour (2007) have shown how children between the ages of two and four use reality observation to infer causal relationships in relatively complex situations). In the domain of reasoning like a scientist, it appears that 6-year-old children distinguish a test that is suitable for discriminating between competing hypotheses from a non-conclusive test when both tests are offered to them for comparison. They are not yet able to imagine a discriminatory test on their own (Sodian, Zaitchik and Carey 1991). However, this capacity evolves with age and it is plausible that in the absence of proper instruction it never attains an expert level. In fact, it is still difficult event for older children (e.g. some studies address 8- to 12-years-old children) to generate by themselves discriminating, unconfounded experiments (Klahr, Fay and Dunbar, 1993; Kuhn et al. 1995; Schauble 1990, 1996; for a review, see: Zimmerman 2000; Morris et al. 2012). Moreover, 8- to 12-years-old children can still confound their initial beliefs with observation data when asked to justify an assertion (Schauble 1990).

As far as the understanding of the nature of evidence and more generally of knowledge (as compared to opinion or to belief) is concerned, research on first-order epistemic capacities in children has crossed its paths with research on metacognitive abilities. E.g., Kuhn (1989, 2005, 2011) has studied the development of children's metacognitive abilities in relation to scientific reasoning on the one hand and scientific reasoning on the other, with the development of CT as a capacity to reason about what constitutes knowledge in regard to a simple opinion. Kuhn was particularly interested in the ability to articulate theorys and empirical evidence, thus initial hypotheses or knowledge and observational data. In her view, this capability involves three sub-components. These include: the ability to represent theory and evidence separately; the ability to treat theories as forms of representation and not as a state of the world; the recognition that a theory can be false and that, to determine whether it is true or false, it is necessary to look to the available evidence. The most important consideration Kuhn arrives at is that strategies for coordinating theory and evidence do not develop without formal instruction. It is therefore not a question of progressive development, in stages, following a natural trajectory, but of learning-driven evolution[97]. We are therefore dealing with skills, which are completely natural, but require cultural integration through education. However, the cited studies focus more on the child's ability to produce than on understanding the value of evidence to support an assertion. As the cited study on understanding a discriminatory test (Schauble 1990) shows, the ability to recognize good evidence may precede the ability to produce it. In the current state of knowledge, further research is therefore necessary to be able to give an opinion on the natural bases of epistemic vigilance with regard to content. (For a thorough discussion of *epistemic cognition*, see Greene, Sandoval, Braten 2016; see also: Carruthers, Stich, Siegal 2009 for research in the naturalization of scientific thinking or the cognitive basis of science).

Before even asking ourselves the question of the evidence supporting a content, or its source, we are spontaneously amnestied to consider its plausibility and relevance. Again, human beings have natural mechanisms for detecting that something is "implausible" to them, and these mechanisms manifest themselves even in the absence of language or without becoming explicit judgments. An exemplary manifestation of such mechanisms is represented by the surprise reaction. Imagine walking into your living room and finding... *a blue gremlin*! You are surprised! You stop at the door and probably for a few seconds you don't really know what to do, especially if you believe your eyes or if you are looking for an explanation in what you have eaten or drunk before. In fact, finding a blue gremlin in your living room is not plausible: it is neither consistent with our past experiences, nor with our knowledge of living things. Surprise is the reaction we experience (and that we manifest outside by signs such as widening eyes, hesitation in our behaviour) when something conflicts with our expectations (Caqati & Pasquinelli 2015).

Many studies of young children are actually based on a fairly simple experimental device that consists of eliciting their reactions of surprise. For example, we get the child used to a certain stimulus and when he has been exposed to it many times we change the stimulus. The child then manifests his surprise with visible signs. This type of paradigm is used with babies as young as a few months old and allows us to explore their implicit expectations - and consequently to highlight the knowledge they have developed early on but cannot express verbally. It is indeed thanks to this type of device that a revolution has taken place in the representation we make of the knowledge of very young children and that we attribute to babies as young as 6 months of age expectations and therefore quite varied knowledge concerning the physical and biological world! What interests us here, however, is not the richness of babies' early knowledge, but the fact that babies - like adults - compare the phenomena they observe with these same expectations and check the agreement between the two; if there is a conflict, a reaction of surprise ensues. We can therefore state that the judgement of plausibility is part of our early natural mechanisms and serves as a basis for our natural CT.

However, as these examples show, plausibility judgement is strongly dependent on our prior experience and knowledge. This consideration has important effects on our strategies for educating critical thinking. We can already anticipate that CT cannot be regarded as independent of knowledge content and therefore educated in an "abstract" way and without enriching the child's knowledge base.

3.2.2 The mechanisms of epistemic vigilance are not infallible

Epistemic vigilance theory emphasizes the idea that the natural criteria we use to assess the quality of information provided are more or less sophisticated and cognitively demanding. In reality, if we wanted to be very selective, we would have to implement more costly criteria that would assess our confidence by adapting to circumstances: for example, by recognizing that the same person may be competent or ignorant depending on the circumstances.

However, we are not always able or willing to pay the cost of these criteria, as a result of which, we use others that are more superficial and have a larger margin of error[98]. Thus, we can base ourselves on the "character" of the person and not on his or her performance under the circumstances - or even on its physical appearance - as indirect indicators of reliability. More generally, we rely on "non-specific" signs. There is therefore a risk that these low-cost criteria may not be refined enough for certain situations where a higher cost may be paid to guarantee against manipulation and incompetence[99].

The performance of the processes of natural epistemic vigilance is deteriorating, especially in complex or unusual situations, in a broader social context - compared to the one in which our abilities have evolved - through social networks for example. So, for example, we tend to judge from the epistemic quality of information if it is widely shared. If each member of the group comes to a conclusion alone, trusting the general opinion is rational. But if it is an idea spread in the group through communication, it can be widely shared for reasons other than its merits. However, because of the situation described above, the index "accepted by many" can be interpreted as a positive signal regarding the epistemic quality of the information. Thus, the very mechanisms of natural epistemic vigilance can become causes of misjudgement, because of the use of of criteria low and at low cost[100].

3.2.3 Social and cultural forms of epistemic vigilance help overcome the limitations of natural epistemic vigilance. Birth of extended epistemic vigilance

Faced with the limits of individual epistemic vigilance, social and cultural forms of this same vigilance have historically emerged. These are institutional bodies made up of experts proposing means for a better assessment of the information circulating within a group. Experts are considered to be those who have more refined and appropriate criteria for evaluating information than the criteria spontaneously used – as compared with non-experts. As in the individual case, collective vigilance must be exercised both with regard to content and source. Science, with its mechanisms for evaluating evidence and its institutional structures for validating sources (*e.g.*, peer reviewing), represents an exemplary case of social epistemic vigilance.

Thus, we can argue that epistemic vigilance is not limited to an individual exercise. As in the individual case, vigilance in this case must be exercised in relation to both content and source.

Science, with its mechanisms for evaluating evidence and institutional structures for validating sources (peer *review* system, reputation, magazines, etc.), is an exemplary case of "social and distributed" epistemic vigilance [101].

We can then consider the articulation between the individual (psychological) mechanisms of epistemic vigilance, the cognitive artefacts (making the evaluation of information more constrained, such as the scientific method) and institutional arrangements for evaluating sources (the academic world, for example, and its system for evaluating the sources of information, comprising journals, *peer reviewing*, and other strategies) allows for an extended form of epistemic vigilance, a form of distributed cognition in the service of epistemic alertness or critical thinking - following a model of distributed cognition (Hutchins 1995a, 1995b[102]).

3.2.4 Explain credulity, irrational opinions and persuasive effects

The existence of natural mechanisms of epistemic cognition and vigilance - and the identification of their limits - have important consequences for any theory of CT.

The first is an optimistic consequence: we are not "**critical idiots**" by nature. On the contrary, we are equipped with protection mechanismss, which are based on the selection of information and its sources. These mechanisms act on the basis of low-cost criteria, are fallible and are particularly challenged when circulating quick remote information. But they are a natural basis for our CT in the face of other people's opinions.

The second consequence is more in terms of questions: how can we explain - in the light of the mechanisms of epistemic vigilance - the dissemination of *fake news* and the "bizarre" beliefs that we (or someone)sometimes hold (e.g. belief in miraculous cures, in a flat Earth, in world-wide conspirations for faking space exploration, and so on)? The rejection of accredited scientific positions (e.g. the rejection of theory of evolution, of climate change, vaccine hesitation)? Does all this only testifies of the limits to our natural scientific cognition and epistemic vigilance? What role does the manipulation of opinion by the media, social networks or particularly influential personalities play in these events?

3.2.4.1 More conservative than gullible

Various social psychology experiments conducted in the 1950s and 1960s by Solomon Asch and Stanley Milgram on the influence of consensus and deference to authority, respectively, left a very negative picture of our ability to resist the influences of others (Asch 1951; Milgram 1974). A widely held view thus sees the human being as "*wired not to seek truth but consensus*": ready to blindly accept the craziest theses because they would be widely accepted or supported by prestigious sources (see Mercier 2017 for a review of this literature[103]).

The theory of epistemic vigilance, and the existence of natural basis for scientific cognition, however, lead to reject the idea of massive gullibility. For Mercier (2017, 2020), this idea is all the more unrealistic since our "default" operating mode is rather conservative. Confronted with information that contradicts our prior beliefs, we engage in a plausibility check. If the new information comes from ourselves (for example, we observe someone in the garden when we thought he was somewhere else), we are ready to change our positios and update our beliefs. But if the information is communicated by others and we do not have good reason to believe in the expertise or good faith of that informant, we tend to camp on our positions. Rejection of the positions communicated by others would be all the more radical as the initial positions and the information received would diverge (Mercier 2017[104], Yaniv 2004[105]). Yaniv (2004) thus tested our susceptibility to accept the opinion of others when we have already formed our opinion on an issue. Even if participants are motivated to give the right answer (they receive a bonus for correct answers), the results indicate that subjects have a biased attitude towards their own opinion. The further apart the opinions are, the less willing the participants are to change their position.

These considerations suggest that the accusations of credulity that are often made and attributed to the harmful influence of others (from the media, television, mass persuasion campaigns, etc.) could be misplaced. A review of the literature by Mercier (2017) suggests that political and religious campaigns are less successful than we imagine, or at least that their impact is modulated by the content of the campaign.

Nevertheless, it seems important, within the framework of CT theory, to quickly explore the issue of the influence of others. In the following paragraphs, we will therefore collect elements to answer at least partially the following question: when do we agree to change our mind and when do we accept second-hand information? Or again: when do we let ourselves be convinced?

3.2.4.2 Good reasons to accept the ideas of others

In the presence of conflicts of opinions, even if we maintain a positive bias in favour of our own opinions, we still engage in assessing the information of others with all available mechanisms of epistemic vigilance, both those of content control (plausibility, relevance) and those of source evaluation (expertise and benevolence). We take into account the informant's skills, possible conflicts of interest, moral reputation, prestige; we pay attention, like children, to his language or accent, thus to his belonging to the same group as us, and to the approval he receives from the group (Mercier 2017[106]). These properties weigh in on our propensity to change our minds and let ourselves be convinced.

Mercier (2017) also insists that, to change our opinion, good arguments must be brought to bear in favour of it. He takes the example of a classic problem of reasoning, that of the bat and the ball, which we present here in a different version [107].

The bottle and its cork together cost 1 euro and 10 cents. The bottle costs 1 euro more than the cork. How much does the cork cost?

This problem is often used as a test of reasoning and to show the existence of irresistible biases that make us irrational (Frederick 2005). A very high percentage of participants fail to give the correct answer (about 85%, depending on the experimental design used, respond that the cap costs 10 cents). Mercier shows, on the basis of experiments conducted online, that the problem is more easily solved in a group setting. Indeed, in a group situation, those with the wrong answer tend to be convinced by those with the right answer, especially if it is presented to them with solid arguments (Trouche, Sander & Mercier 2014; Trouche, Shao & Mercier 2019; Mercier 2017). Our propensity to follow the opinion of others would thus be modulated by the mechanisms of epistemic vigilance already highlighted (with their limits) and by the existence of good reasons for doing so.

3.2.4.3 The influence of content in the acceptance of ideas

However, we do not rely exclusively on good reasons and arguments to follow the opinion of others. We have cited the literature review by (Mercier 2017) that the influence of politicians, religious and advertisers on our opinions is less than we imagine. This influence does exist, however, modulated by the content promoted[108]. Thus, we can let ourselves be convinced that our intuitions are not in conflict, but rather in agreement, with the ideas promoted by others, be they good or bad. (An example of how our own intuitions can induce us to accept false ideas is represented by the practice of bleeding: totally ineffective, dangerous, and yet widespread. Some attribute its success to Galen's teachings, but was actually rooted in our intuitions of purity[109]).

The idea that certain characteristics of an idea make it more attractive than others dates back to a long tradition: to the idea made popular by Dawkins (2016, 1982[110]), to the theory of the epidemiology of representations developed by Cavalli-Sforza and Feldman (1981) and by Sperber (1996) and to the theory of cultural transmission proposed by Boyd and Richerson (2005), Richerson and Boyd (2008). All three present the application of a Darwinist approach to the evolution of culture, through the idea that cultural information is disseminated in a population through its interactions and the production of social artifacts[111]. Within this tradition, the influence of emotions - including strong emotions (Heath, Bell & Sternberg 2001), negative emotions (Bebbington et al. 2015) and those associated with danger Blaine and Boyer (2018) - has been used to explain the success of urban legends (Heath, Bell & Sternberg 2001), religious beliefs, rumors (Blaine & Boyer 2017). More recently, Acerbi (2019a, b) analyzed sites known to publish *fake news*. He thus highlighted that the information put forward had specific characteristics compared to that considered less attractive. In particular, *fake news* more often involved negative threat-related emotions, contained sexual elements, and appealed to reactions of disgust[112]. Thus, according to these examples, our epistemic vigilance type protection systems are not without gaps and imperfections (Boyer 2018[113]). Not only are these systems limited[114], but the characteristics of certain information content create a gap in our defence systems because their content is particularly salient, attractive, or adaptively relevant[115]. These contents therefore make it even more difficult to exercise vigilance.

To these considerations, it is necessary to add the fact that the evolution of our vigilance systems is leading to the development of even more powerful weapons of persuasion and manipulation. Since deception is advantageous, systems to deceive evolve along with those to counter deception. The design of these systems could therefore, at least for certain contents and under certain conditions-or at a specific point in the arms race-give an advantage to of strategies of deception[116].

3.2.4.4 Social-driven mechanisms. The role of the group

Other reasons for the success of successful ideas may seem less immediate. These include the ability of some ideas to cement alliances. According to Atran (2002) and Boyer (2018), some of them are particularly suitable for the creation of groups with a very high level of commitment. These include counter-intuitive ideas, such as the existence of ghosts, miracles, etc., and the need for a group to be able to understand and act on them. Adhering to counter-intuitive ideas would be a strong signal of belonging to a group. The very adherence to these ideas would therefore be motivated - implicitly, not assumed - by the desire to be a member of the group at all costs. Similarly, some rumours would actually be used as consensus builders. Whoever passes on the information has not (at least not only) the motivation to inform but (above all) the motivation to create a coalition for joint action (Boyer 2018[117]). This type of mechanism could partly explain the success of conspiracy theories and of the most counterintuitive religious beliefs, as well as the most unbelievable *fake news*.
A stream of studies in social psychology has developed a cognitive and cultural theory to explain certain anti-scientific positions - including anti-vaccine and climate change sceptic positions (Kahan 2015). This theory helps to explain resistance to cognitive explanations when adherence to ideas involves membership in a social group, and when abandonment of these ideas involves the disintegration of the group itself.
Should we consider all these reactions to be irrational? It all depends on what we mean by the term. Information about dangers and threats is certainly useful from a survival perspective. It is not surprising then that we are particularly inclined to listen to information with this kind of content, whether it is true or false. Group membership is certainly a condition for survival for a social and cultural species like ours.
Our natural equipment - the fruit of our long evolution - is formed by the selective pressure to collect information useful to our survival rather than worrying about filtering scientifically

proven knowledge[118]. In a sense, adaptive, it is therefore rational to give priority to certain contents, even if they may turn out to be false in the light of rigorously established facts.

3.2.4.5 Social-driven mechanims (2). The role of the prestige of the source of information

If there are elements in the content of the information that can mislead our natural defenses, This could also be true for certain characteristics of the information sources. One stream of studies focuses on the mechanisms of social learning, which is considered very important in explaining the accumulation of culturally produced knowledge (Chudek & Heinrich 2011). More generally, like communication, social learning is subject to rules. Thus, the choice of "masters" or examples to follow and imitate in order to learn involves mechanisms close to those of epistemic vigilance, which determine the evaluation of the quality of information. Robert Boyd, Peter Richerson, Joseph Heinrich and other cognitive anthropologists have studied the criteria and biases that influence social learning., The model's characteristics, such as age, gender, health status, language, ethnicity, or model expertise (Boyd & Richerson 2005; Richerson & Boyd 2008; Chudek & Heinrich 2011). By bias, we mean here a criterion that influences a choice, even unconsciously. As with epistemic vigilance, relying on often indirect and low-cost criteria can be a potential source of error. This is particularly evident in the case of expertise, which is difficult to evaluate directly and therefore inferred from clues more or less related to the real capacities of the model, such as: perceptible signs of competence, signs of previous success, signs of prestige, which result in the model being followed and copied by others[119]. This last criterion is considered a particularly risky index. By virtue of this bias, an individual can be followed and imitated - and thus become influential - independently of his/her actual skills in the field of expertise (Henrich & Gil-White 2001).

Chudek et al (2011) tested the existence of prestige bias in children. They exposed children of about three years of age to two models. One is a prestigious model according to the definition presented above: it has followers, assistants of the experimenter who look at it attentively. The other does not receive this type of attention from the assistants and is therefore considered non-prestigious. Both models, prestigious and non-prestigious, perform different tasks: they name objects, express their food preferences for different foods or objects. The results indicate that the decline in prestige is present in different areas of cultural learning (including food preferences) but its strength varies from one area to another[120]. A recent review of the literature suggests, however, that the influence of prestige bias in social learning is reduced and modulated by factors such as the difficulty of the task to be learned, the importance of the task and the presence of other social learning biases - particularly content-related biases[121]. How the characteristics of the information source can attack our natural defences is therefore still to be explored.

We know that remote communication changes the way of deploying and therefore detecting signs of expertise. Some visible signs can be added (*likes*, votes, followers) and possibly falsified. It is therefore crucial to understand how these signs are produced and what they actually translate in these new communication and information sharing environments.

3.2.4.6 Uncertainty as a position changer

Several studies show that confidence in one's knowledge (an aspect of metacognition) is a fundamental modulator of the ability to change one's position and the adoption of other people's positions against one's own (Mercier 2017).

For example, Yaniv's (2004) study on readiness to change positions and take others' opinions, cited above, shows that participants take into account their level of knowledge and use this self-assessment to decide whether to take or reject others' opinions: the more they think they know, the less they listen to others' opinions; the less knowledgeable they think they are, the more willing they are to change their first opinion after listening to others' opinions. We have seen that the further apart the opinions of others are, the less they influence the final opinion, but, again, this depends on the confidence that the subject places in his or her own knowledge[122].

The influence of uncertainty on the susceptibility to accept the opinion of others has already been shownce in the behaviour of eighteen-month-old infants who are faced with deciding whether or not to go through a difficult passage (Tamis-LeMonda et al. 2008). In a study conducted on about twenty pairs of subjects (mother and child), Tamis-LeMonda et al. (2008) observed that children tend to ignore mothers' advice to cross a slope if the slope is perceived as too dangerous; they also ignore mothers' advice not to cross if the slope is perceived as easy. In the intermediate condition, a slope that is neither too easy nor too difficult, the child's behaviour changes according to maternal encouragement, with more likely to cross if the mother encourages than if she discourages[123]. We are therefore dealing with an early sensitivity to the other's opinion in relation to the degree of confidence the child has in himself and in his perception of danger.

The role of uncertainty has also been explored in studies of social learning and cultural transmission in other animals[124]. The problem is to determine when an individual has an interest in imitating another (social learning) rather than seeking to learn by himself (individual learning). Like any other form of communication, social learning is indeed advantageous only if certain conditions are given, such as the reliability and expertise of the model to be imitated. This is as true for our species as it is for other social species. Using models, researchers have shown that a fundamental condition for social learning in different animal species is uncertainty. They have called this condition the "copy bias if uncertain". This bias has been demonstrated not only using mathematical models but also empirical studies on non-human species (Rattus *norvegicus*: Galef, Whiskin & Dudley 2008; Pungitus *pungitus:* Van Bergen Coolen & Laland 2004).

This gives a strong indication to look for other cognitive building blocks of CT in the metacognitive capacities related to the perception of oneself, either as competent and performing or as uncertain, and to the perception or anticipation of error in one's choices and decisions.

Before reviewing the relevant literature, let us conclude this first part by presenting the implications of the themes discussed for CT education.

3.2.6 Epistemic cognition and Epistemic vigilance. Consequences for education

Mechanisms called “cognitive epistemology” and "epistemic vigilance" allow us to use our observational skills and our soocial skills of communication, and transform them into useful sources of knowledge and information. They create a discriminatory filter that operates on the basis of plausibility, relevance, supporting evidence in order to assess the truthfulness of contents and they screen our sources on the basis of their trustworthiness - whether they have reasons to cheat and manipulate us, because of self-serving interests, or because of their (lack of) moreal values, whether they are sufficiently reliability, based on their general knowledge and specific expertise.

These mechanisms therefore form the basis of our ability to assess the epistemic quality of information in order to make a decision - a capability we have identified with the CT.

For this reason, **we consider epistemic cognition and epistemic vigilance mechanisms to be fundamental to CT**.

Moreover, **since these mechanisms are natural, it reinforces our view that CT is - also - natural.** We have, however, highlighted the **limitations of naïve epistemic cognition and epistemic vigilance and, therefore, CT.** We have also pointed out that these mechanisms are likely to be misleading because of intrisic liitations, the indirect nature of the criteria used to assess content and sources of information, and competing interests.

Henceforth, breaches in our natural defences are quickly made. This is all the more true since the selective pressure that allowed the development of the mechanisms of epistemic vigilance probably took place in contexts quite different from the one in which we are currently living, particularly in terms of information exchange and sharing systems.

The goal of CT education is to maximize the ability to assess the quality of information for decision-making in a variety of situations, both current and complex.

Given the existence of natural mechanisms for evaluating information, developing CT therefore essentially means equipping it: **exercising more sophisticated criteria than those used spontaneously to respond to more complex situations in a correct way**.
It is therefore necessary to take into account the natural capacities and criteria spontaneously used for the evaluation of information sources and contents, and their limitations.
It is not about creating capabilities *ex novo*, nor is it about developing reasoning, rationality or thinking in general. It is much more about providing knowledge - about what constitutes a reliable source in a certain context, for example, or what constitutes solid evidence to support a claim.

The concept of context-specific CT helps to distinguish an advanced CT from an expert CT. The advanced CT is CT that is equipped to respond to common day-to-day challenges. In theory, everyone in everyday situations has knowledge specific to advanced CT. The expert CT is based on specialized knowledge, which is not useful for everyone in all circumstances, but which allows for the proper evaluation of information in specific contexts.

The main indication for education is to provide learners with **knowledge and criteria for evaluating information**:

- to assess the relevance of information provided in support of an argument;
- to assess the plausibility of the information content;
- to assess the validity of the same information content on the basis of the available evidence;
- to assess the expertise of the source of the information;
- to assess her benevolence (her hidden interests and anything that might justify or make us think that she wants to deceive us).

These acquired tools are then integrated into natural tools, but they are more sophisticated than those used spontaneously for the evaluation of contents and sources, and better adapted to the present context.

The assessment of the relevance and plausibility of information depends, among other things, on whether or not one is knowledgeable in the field concerned. An idea appears plausible because of the knowledge base available. For example, the sentence "This crystal has the power to heal through its energy" may be considered perfectly plausible by someone who is unfamiliar with the scientific concept of energy. The idea that homeopathic pills can cure diseases is plausible if we ignore their actual composition or the fact that water memory is not a scientifically validated theory.

From this consideration flows the importance of **not dissociating CT education from the acquisition of factual knowledge**, i.e., scientifically validated knowledge about the physical, biological, social, etc. world.

Assessing content validity involves the ability to use observations and supporting facts or evidence correctly. This is a difficult task, which is why, in the case of second-hand information, the assessment of sources often takes precedence over the assessment of content (except for the assessment of relevance and plausibility, i.e. consistency with one's own knowledge). Suppose we wanted to evaluate a statement such as "This homeopathic vaccine prevents influenza" on the basis of the supporting evidence. To do this, we would need to know that individual observations (anecdotes) have very low evidentiary value. In addition, we would need to know the added value, to avoid observational bias and confounding factors, of experimental testing with control groups and randomization, and repetition of the same tests. With this knowledge, it would be possible to compare the evidentiary value of testimony provided by a relative with that of evidence from scientific studies using the methods cited. It is precisely this type of knowledge that can enrich our natural CT and make us better able to distinguish informational content on the basis of the quality of supporting evidence.

This knowledge also provides an additional objective. Faced with the difficulty of practicing this type of content evaluation on a daily basis, understanding which methods provide the most certain evidence allows for the identification of more reliable classes of information sources on the basis of their expertise. Scientists, who apply rigorous methods to establish cause and effect relationships, deserve our trust by virtue of this same practice. Knowing methods for validating knowledge (from rigorous observation to experimentation, peer review and other strategies for social control over individual outcomes) allows us to better identify not only stronger content but also more reliable sources.

- From this consideration stems the need to exploit scientific lessons in order to provide an increasingly thorough knowledge of strategies and methods to control the main observation biases and to properly articulate theories and observed facts. It is not necessary - in order to demonstrate CT - to be able to set up an experiment or conduct rigorous observation in the manner of a scientist. Knowing these strategies and methods, their reasons and consequences, however, makes it possible to distinguish between information content based on evidence of non-validated content, and reliable sources by virtue of the methods used to reach a conclusion from less reliable sources of expertise.
- A certain amount of scientific literature, relating to the methods and workings of science, makes it possible, among other things, to recognise the specific nature of scientific knowledge and to identify science as a privileged source of knowledge concerning natural phenomena.

Assessing the validity of content involves the ability to correctly use observations and supporting facts or evidence.

3.3 The cognitive bases of CT: metacognitive sensitivity, confidence

Let's go back to the example of the restaurant. This time, instead of asking someone's opinion, we choose to use our observations as the basis for choosing "the best restaurant". We look at the line in front of the restaurant, its appearance, the look of the food. Based on this first-hand information, we make our decision.

In both cases, first- and second-hand informatio the question of *trust* emerges. In the case of second-hand information, it is the *trust* we have in our sources or in the content of the information that is delivered to us, according to the evaluation by criteria that we have just described. In the case of first-hand information, which we collect using our senses or which is stored in our memory, it is the trust we have in our own ideas, in our perceptions, in our memories and in ourselves as sources: "Should we feel sure of our opinion, of our decision? Should we doubt? Do we feel uncertainty?"

One need only be in a dark and unfamiliar place to experience an immediate drop in confidence in the information provided to us by our sensory organs. Our steps become more uncertain, we slow down our walk to accumulate more information as we go along and thus eliminate the hypothesis of possible obstacles in our way. We know that it is easier to make mistakes under these conditions. We don't even need to consult with ourselves to know that our feelings of uncertainty are affected by the situation. Even without this becoming explicit, our decisions are affected by our feelings of confidence in the information available and our chances of getting through the situation without breaking.

This example of walking in the dark is one of many where information assessment, decision-making and trust are silently but securely linked.

How do we know when it's reasonable to trust our decisions, feelings, opinions? How do we manage to estimate our chances of being wrong, our chances of making a mistake rather than making the right decision? Does our high level of confidence mean that we do not make mistakes, whereas low confidence means that we are more likely to make mistakes?

Note that, in the following pages, we speak about confidence in a certain perception, knowledge, opinion, belief and decision, i.e., confidence in a certain representation, not self-

confidence or confidence in our decisions, in general. We want to know how confidence in these mental states influences decision-making and learning (by decision, we mean a mental state of making a choice when there are several options available. A decision can therefore be purely perceptual in nature: faced with an ambiguous sound stimulus, we decide that it is a rain noise; faced with an uncertain visual stimulus, perceived in a dark environment, we decide that it is a cat. Learning is here defined as the updating of predictions, representations, prior decisions in the light of new evidence - evidence, feedback, new observations, new stimuli. It is therefore related to the examples of changing positions and opinions - in situations of uncertainty and/or in the light of evidence from outside).
The literature on confidence has been rapidly expanding over the past decade or so. This literature mobilises a wide range of disciplines from the family of cognitive sciences: philosophy, neurosciences - neuro-imaging, modelling -, developmental psychology, experimental psychology, reasoning and social psychology (Proust 2013; Fleming & Lau 2014; Grimaldi, Lau & Basso. 2015; Goupil & Kouider 2019). It is situated in a broader framework: that of studies concerning metacognition and, in particular, implicit metacognition and metacognitive sensitivity.

3.3.1 Basic aspects of metacognition

Can we trust this information? Is it reasonable to use this information to make a decision? Is the decision we have just made justified by the information available? Asked explicitly, to ourselves, these questions refer to a particular group of mental states managed by metacognitive processes. A classic definition of metacognition is "*thoughts about one's own thoughts and cognition*" (Flavell 1979). Thus defined, metacognition appears to be a form of second-order cognition (cognitive states having as their object other cognitive states of the same individual) (Grimaldi, Lau & Basso 2015). The functional role of metacognition is said to be twofold: that of monitoring mental states (Dunlosky & Metcalfe 2009) and that of controlling these states. Metacognition is thus defined as the ability to monitor and control cognition, or the ability to reflect on our mental representations in order to regulate cognition and optimize learning. It includes several processes such as the regulation of attention, the choice to be trained or to be informed, and the action of identifying and correcting one's errors (Proust 2019[125]).

The fields of application of metacognition are vast and varied and, as a result, meanings of the term may vary slightly from one author to another (Fleming, Dolan, Frith 2012[126]). Even before metacognition became a research topic in itself in psychology, notably thanks to Flavell, studies on memory have explored metacognitive experiences such as the feeling of having a word on the *tip of the tongue* (ToT) and the *feeling of knowing* (FOK). Metamemory is thus a fairly well-established field of research. In the meantime, however, studies have been extended beyond the boundaries of psychology. Metacognition is now a multidisciplinary research subject, particularly represented in the fields of educational psychology and cognitive psychology, but also in clinical, animal and neuroscience psychology (Norman et al. 2019). Among the most explored functions related to metacognition are learning (e.g., Bransford, Brown, Cocking 2000), the relationship of knowledge to belief (e.g., Wellman, Cross & Watson 2001), and decision such as, for example, the perceptual decision: "Is the object I see in the dark really a cat? What object makes that noise? I feel like I can hear the rain falling, but in reality it is not raining" (Norman et al. 2019). In decision-making situations, metacognition intervenes at two levels: to anticipate the chances of error and to gauge the reliability (the degree of certainty) of a decision/information in relation to others. By virtue of these components, metacognition makes it possible to guide our behaviour in complex situations: for example, in the presence of multiple sources of information or ambiguous information that lend themselves to different interpretations. Thus, correctly estimating the trustworthiness of a certain piece of information allows us to decide whether it is necessary to seek out new information or whether the available information is sufficient. We then speak of "metacognitive trust" and define it as a probability judgment about an action or the likelihood that a given judgment is correct. Metacognitive trust could also have a functional role at the level of social cognition, for example in the case of collective decision-making, in order to resolve possible conflicts of opinion in favour of more reliable ones, or to enable a change of opinion in favour of a more reliable source (Fleming, Dolan & Frith 2012). (In the previous section, we anticipated, based on knowledge from behavioural studies, that uncertainty was a driver for changing one's position, i.e., updating one's opinions and judgements to adopt those of others).

3.3.2 Implicit metacognition

Metacognition appears in Flavell's definition as an intentional state in which we voluntarily take our mental states into account in a reflexive way.

However, the concept of metacognition is broader than that of explicit or declarative metacognition (with verbal transmission of beliefs, naïve or expert theory about cognition) or conceptual metacognition. Classically, we distinguish metacognitive knowledge (I know what cognitive state I am in, I understand my cognitive abilities, I know what allows me to improve them), metacognitive strategies (I implement actions that will allow me to improve my performance, or at least to control it), metacognitive experiences or feelings (I feel I am close to the goal, or in difficulty in a task, I have a feeling of familiarity with an exercise, the feeling of being able to succeed) (Flavell 1979; Efklides 2008, 2011). The metacognitive feelings produced during the activity have an important role in predicting the failure or success of a task, and thus in the metacognitive self-evaluation that precedes the response to the task itself (Proust 2019, Koriat 1993[127]). These feelings emerge as the visible tip of the iceberg from processes that take place silently in our brain.

The brain uses predictive cues to estimate its chances of success in a task - to estimate its confidence in its ability to solve a problem, make the right decision, its chances of responding correctly or making a mistake. These cues are indirect: they are "predictive heuristics". They can be the speed with which information related to the task is processed (fluence heuristic) or the consistency between the response produced and other representations activated by the trace (consistency heuristic). At the neural level, we identify as a mechanism of trust the convergence between different groups of neurons towards a single decision (magnitude of activation produced by a task) (Proust 2019[128]). Metacognitive feelings thus belong to a form of procedural, implicit, sub-personal, silent metacognition - which does not need words to express itself and act on behaviour.

3.3.3 Developmental trajectories of metacognition

Explicit, verbalized metacognition matures gradually during childhood. For example, the child is able to express correctly what he or she knows or does not know around the age of seven or eight. His ability to estimate the difficulty of a cognitive task improves in the early years of school. The sense of knowing and the ability to judge his or her learning also improves, but to a lesser extent. The development of strategies appropriate to the difficulty of the task occupies the latter part of the primary school years.

Nevertheless, even much younger children show themselves capable of estimating [129]their level of knowledge - implicitly - and seeking additional information - or refraining from giving an answer if they are uncertain (see below). Estimates of certainty and uncertainty are therefore present in some form, regardless of whether they can be expressed verbally. We find evidence of implicit metacognition in adults, as well as in young preverbal children and non-human animals (Palmer, David & Fleming 2014; Smith, Shields & Washburn 2003; Grimaldi, Lau & Basso 2015). Their expression is behavioural and is highlighted by silent experimental devices such as:

- the opportunity given to the subject to seek out other information in addition to that in his or her possession;
- the possibility of not making a choice between two options, but to opt out of the experiment (*opt-out)* so as to minimise losses while giving up the possibility of winning a greater reward: the subject, prospectively, decides in advance not to give a certain type of answer if he knows that the reward is rather associated with a correct answer (large reward) and no answer (small reward), but loses the chance of a reward if he gives an incorrect answer.

The idea being that the subject estimates his or her level of knowledge before responding and acts accordingly to maximize benefits) (Norman et al. 2019; Shea et al. 2014). These devices help overcome the prejudice that metacognition is limited to those organisms that are proficient in language. Their nature, however, leads to different interpretations of the results obtained - especially when the tests are carried out on animals - and keeps open the debate on the relationship of metacognition to consciousness. For example, animal studies of trust are often criticized on methodological grounds. Indeed, it is difficult to rule out the possibility that mechanisms other than the evaluation of trust are at stake. It is, however, fairly widely accepted that non-human beings (animals) are capable of estimating their confidence in a certain decision or knowledge and, more generally, that they possess implicit metacognitive mechanisms (Fleming, Dolan & Frith 2012[130], Shea et al. 2014).

Some authors have proposed to distinguish between automatic, implicit, index-based metacognitive processes on the one hand, and more demanding, supra-personal metacognitive processes on the other (Shea et al. 2014). The former allow to control neural activity even at very low levels of processing, taking into account the uncertainty relative to each representation, and to control the activity of different perceptual and motor systems. The latter allow this confidence to be expressed explicitly, even verbally. They therefore have a particularly important functional role in cooperation between individuals, when it comes to taking decisions with several people, taking into account the uncertainty and reliability of each one. Other authors - as we shall see later - have proposed the introduction of neural processes that read information about trust from the properties of the activity of the neurons involved in a decision (Meyniel et al. 2015). In both cases, what is important to remember concerns the omnipresence of implicit metacognitive processes and their dependence on indirect cues, and the idea that other processes can come into play by exploiting implicit processes and ultimately improving them.

3.3.3.1 Early mechanisms of metacognitive sensitivity

Goupil and Kouider (2019) provide a synthesis of current knowledge regarding implicit metacognition in young children. This synthesis shows that children from three to five years of age are able to express well-calibrated confidence in their decisions through non-verbal means, and that children from twelve to twenty months of age display implicit metacognitive abilities. These findings contradict a well-established view of metacognitive development that this function is rare in childhood (Flavell 1999, Schneider 2008, Sodian et al. 2012).

Children between the ages of three and five can be trained to use a non-verbal confidence scale (images of a confident or doubting child). Using this type of device, they can then express their confidence in the information in their possession (retrospective, explicit confidence). We find that correct answers are correlated with greater confidence and incorrect answers with less confidence: children of this age therefore show that they are capable of correctly assessing their performance and have metacognitive sensitivity.

The notion of metacognitive sensitivity is a normative notion: a subject shows metacognitive sensitivity when he is confident in the decision [131]he has made or to be made and that this confidence is associated with effectively correct performance. Similarly, metacognitive confidence is well-calibrated when confidence is low and performance is unsatisfactory. Conversely, in cases of poorly calibrated trust, trust and reality diverge (Fleming & Lau 2014).

A classic distinction within metacognitive trust is between retrospective and prospective metacognition (Grimaldi, Lau & Basso 2015). Retrospective metacognition is measured in the following way: subjects who participate in the experiment are asked to evaluate their success in a task i whether the task has been completed[132]. The second is measured by proposing a task and asking subjects to anticipate their chances of success; in this case, subjects can decide whether to respond or not (as in the case of the *opt-out* paradigm described above)[133]. Foresight-type paradigms are therefore more often used to study implicit metacognition.

Opt-out paradigms have shown that the ability to estimate the probability of making a mistake is already present in the twenty-month-old child and that metacognitive sensitivity to the correctness of decisions is present as early as twelve months of age. The following are a few examples.

Twenty-month-old children need to remember where a toy has been hidden. To vary the difficulty, a more or less long time is given between the time the toy is hidden and the time the child has to find it. Children in the control group are forced to make a choice, while children in the experimental group have the option of *opting out* by asking an adult for help. The rate of correct responses compared to the use of adult assistance indicates that children correctly estimate the risk of making a mistake, and thus their uncertainty (Goupil, Romand-Monnier Kouider 2016[134]).

To calibrate confidence in the decision made, Goupil and Kouider (2016) used the same paradigm but with one variation: measure persistence in finding an object hidden in a box according to the degree of certainty that the object is actually in the box[135]. The authors conducted their study with eighteen-month-old children, who observed an adult hiding an object in a box. The delay between the time the object was hidden and the time the child had to point with his or her finger to which box the object was hidden varied. According to the authors, uncertainty increases with increasing delay because performance (finding the object hidden in one box rather than another) decreases with time. They then proposed the hypothesis that persistence in the search for the object hidden in the box is all the stronger when the child is sure to find the object in the box he has indicated, and is therefore confident in his choice. They have found that, indeed, children search longer - persist more - if their answer is correct. They interpreted this as a sign of correct calibration of metacognitive confidence in decision-making (Goupil & Kouider 2016). Another test of persistence involves hiding the desirable object (a cookie) either in an easy-to-open box or in a box with a sealed lid, which requires adult help to open. Similarly, the time delay between the time the object is hidden and the time it must be found varies (three to twelve seconds). If the object is hidden in the sealed box, the child - convinced that the object is inside - must ask for the adult's help to open it. If they are not really convinced, they try to open the other box to see. The persistence in his choice therefore depends both on having made the correct choice and on the time elapsed (Goupil & Kouider 2016).

The same type of ability was measured in children aged five to eight years using the so-called "post-decision betting" paradigm. The children observe pictures containing objects (dots) in varying amounts. Following the estimation ("Which of the two pictures contains the most objects?"), they must express confidence in their judgment ("How sure are you of your answer?"). The expression of confidence involves choosing between an emoticon representing a smiling face and another representing an uncertain face. The results show a positive correlation between task difficulty and performance, and between the latter and expression of confidence (metacognitive sensitivity). In this test, metacognitive sensitivity improves with age from five to eight years (Vo et al. 2014).

Finally, further evidence confirming the presence of an early form of metacognitive sensitivity in children is provided by neuroimaging in adults and in other species, such as rats and monkeys (Goupil & Kouider 2019). In adults, metacognition involves the activation of the prefrontal cortex: confidence in decision-making seems to be encoded in the ventromedial prefrontal cortex; error monitoring (lag between the erroneous choice just made and the choice that should have been made on the basis of the available evidence) is represented in the anterior cingulate cortex and by a characteristic signal in EEG (ERN: *Error-Related Negativity)* (Goupil & Kouider 2016). These structures are already active during childhood (e.g., the ERN component is present in EEG after incorrect selection in 12-month-old children), confirming that metacognitive sensitivity is not limited to explicit abilities in adults. The authors therefore suggest the existence of an early metacognitive system, or *core metacognition*, which would be present in children before they are capable of verbal intercourse and also in other species. The function of this early system would be to evaluate and regulate cognition in an automatic and implicit way, in particular to evaluate the quality of a representation: the probability (*likelihood*) that the representation is correct, given the available evidence (of a perceptual type). Based on the fact that the prefrontal cortex matures more slowly than other cerebral regions such as those involved in perceptual decision-making, we can assume that the maturation of the core metacognition is also slow. The core metacognition system would also have other limitations of its own: the fact that assessment involves the use of cues, such as the time taken to make a decision, rather than more reliable calculations. The fact that implicit evaluations cannot be transmitted verbally makes it impossible to exploit core metacognition in a social exchange to communicate its level of certainty. To achieve this, core metacognition needs maturation and the entry into play of other abilities, such as the ability to focus attention on metacognitive representations, making them explicit or conscious. Thus, we are in the presence of a problem of attention and of the salience of the representation in relation to external stimuli. We can also assume that the overconfident attitude of children (and adults) who systematically answer "Yes, I know" is linked to the inability to select (selective attention) the right source of information: the error signal rather than the desire to succeed. This would therefore be a defect in executive capacity. The authors question the role of culture and social interactions on the development of metacognitive systems (learning to express metacognitive states, for example).

3.3.4 Confidence and the brain

The notion of trust has recently crossed the barrier of psychology to interest the neurosciences. This interest is reflected in an effort to identify brain regions and, more generally, neural substrates that correlate with observable metacognitive trust behaviour[136]. This effort involves studies in neuropsychology - using TMS, structural MRI, fMRI, selective pharmacological inactivation. There is no consensus for the moment on how trust would be encoded at the neural level: either *via* dedicated modules separate from those for decision-making (possibly with multiple brain representations and partly different because of the content), or rather *via* circuits dedicated to decision-making (Grimaldi, Lau & Basso 2015). The hypothesis of specialized networks seems dominant in the literature, particularly for its ability to explain cases of disjunction between performance (decision) and confidence estimation (Grimaldi, Lau & Basso 2015; Meyniel, Sigman & Mainen. 2015).
Neuroscience has also focused on defining trust in purely neural and computational terms. This definition includes the notion of the Bayesian brain and models of the brain as an

essentially probabilistic organ (Kepecs & Mainen 2012; Meyniel, Sigman & Mainen 2015; Grimaldi, Lau & Basso 2015).

3.3.4.1 Confidence and the Bayesian brain

The Bayesian Brain Hypothesis is now held in high regard in order to explain how the brain represents reality from relatively poor, uncertain or ambiguous stimuli (Pouget et al. 2013[137]). Or how it makes decisions (including perceptual ones) and how it updates them based on prior knowledge (learning). Bayesian models make it possible to account for a variety of observed phenomena, such as perceptual illusions and biases, but also the baby's statistical learning modalities (Meyniel 2016). These models are based on the Bayes rule: a mathematical tool introduced by the Reverend Bayes that allows us to express the probability that a certain hypothesis is true according to the available data, the *a priori* plausibility of the hypothesis itself (in relation to other knowledge) and the probability of observing the present data if the hypothesis is true[138]. According to the Bayesian Brain Hypothesis, every representation that the brain makes of the environment is therefore in reality expressed in probabilistic terms and takes into account not only the present stimulus but also a priori factors that influence the way the stimulus is interpreted.

Let us now see how this model allows the notion of confidence to be integrated into the very heart of brain functioning[139]. Let's take the example of a perceptual decision where the brain is faced with a slanted line: by how much is the line slanted? The fundamental aspect of the brain that is important to consider here is that a population of neurons - not a single neuron - responds to the stimulus (the slant line). However, each neuron that activates in response to the line does so in its own unique way. Thus, each neuron responds in a privileged way to a certain inclination of the line: for example, it responds with a stronger activation if the line has an inclination of 45°. Other neurons respond to lines oriented at 45° but in a weaker way. Still other neurons do not respond to lines oriented at 45°, but only to other orientations. The overall distribution of activations in the neuron population then corresponds to an estimate of the probability that the present stimulus has a certain orientation - for example 45° tilt (Knill & Pouget 2004; Meyniel, Sigman & Mainen 2015). The perceptual decision (the line is oriented at 45°) is a probabilistic decision. If we look at the variance that exists within this distribution, we also get an implicit measure of the uncertainty that accompanies the estimation of this probability: the higher the variance, the greater the uncertainty. Uncertainty is considered here as the flip side of confidence, so we can say that the brain's response to the stimulus also contains an implicit assessment of uncertainty *and* confidence (Meyniel, Sigman & Mainen 2015). Pouget, Drugowitsch and Kepecs (2016) rather define confidence as the estimate of the probability that the perceptual choice made is correct, based on the available evidence (stimuli), but distinguish it from the notions of certainty and uncertainty, which do not necessarily imply a choice[140]. However, this difference in definition is not significant here. There are indeed other models, still of Bayesian nature, that allow us to represent a perceptual decision at the brain level and to define confidence at the neural level. We limit ourselves to the model discussed, as the aim is not to provide a complete picture of neural approaches to confidence (for this, we can consult: Dunlosky & Metcalfe 2008; Kepecs & Mainen 2012; Middlebrooks & Sommer 2012; Shadlen & Kiani 2013; Fleming & Frith 2014; Fleming & Lau 2014; Grimaldi, Lau & Basso 2015; Meyniel, Sigman & Mainen 2015; Pouget et al. 2016), but to present the general principle that it is possible to represent confidence at the neural - subpersonal - level as a consequence of the state of activation of a population of neurons, within probabilistic models of brain functioning. A psychological concept, that of confidence, is thus translated in purely neural, and purely implicit, terms (see also, in this respect, Ott, Masset & Kepecs 2018).

Meyniel and colleagues (2015) use the term "*distributional confidence*" to present this way of conceiving trust at the neural level. Distributional confidence is inseparable from the decision. In our example, the uncertainty related to the interpretation of the orientation of the line is indissociable with the probability that the line has a certain inclination. One question is whether this type of description can be applied beyond the specific case of the perceptual decision to judgements of metacognitive confidence at a higher computational level, such as those described in the psychological literature (Meyniel, Sigman & Mainen 2015; Ott, Masset & Kepecs 2018[141]).

Another question, which we anticipated above by evoking the specialized circuits for metacognitive confidence, concerns the possibility that other processes and mechanisms exist and carry out a reading of certain properties of the distribution of activation (Meyniel, Sigman & Mainen 2015). These processes and mechanisms are grouped under the term "*read out"* or "*confidence extraction*". For example, we can hypothesize the existence of mechanisms that separately link the direction to which the activation distribution points and the variance in the activation distribution (or other mechanisms).

Postulating the existence of *read out* processes makes it possible to obtain - from a similar distribution of activations within the same population of neurons - a separate and specific representation of confidence in relation to the representation of the decision ("*summative confidence*" vs. "*distributional confidence*", according to the terminology used by Meyniel, Sigman & Mainen 2015). The two processes are both united (because they take place in the same neuron population) and disjoint (because they are read by two different mechanisms[142]). The *read-out* hypothesis thus makes it possible to explain certain results obtained in neuroscience according to which, at the level of brain activation, the circuits of decision-making and confidence are partly, but not totally, superimposable (Meyniel, Sigman & Mainen 2015[143]).

According to Meyniel, Sigman and Mainen (2015), *read-out* is then the kind of process by which it becomes possible to provide a single, scalar estimate of the confidence we have in a decision, perception or opinion. When, in a verbal or explicit test, a subject expresses his or her level of confidence, it refers to the confidence *read-out* result that accompanies the decision. When he or she decides to provide or not to provide an answer in an opt-out or betting test paradigm, he or she implicitly refers to the same single *read out*.

There is no obligation to *read out* once information is probabilistically coded. Trust can remain encapsulated within the circuit that processes the information in question. But, if this *read out* is performed, it becomes possible to express and thus share and compare summative trust. This sharing can take place within the same cognitive architecture - of a brain, but also between several regions that code for different types of stimuli or sensory and other modalities. Sharing also becomes possible between individuals. *Read-out* processes could therefore play a functional role in cooperative situations, to reach an optimal collective decision taking into account the uncertainties of each individual and, more generally, in all complex situations where multiple sources of information and knowledge are to be taken into account (Meyniel, Sigman & Mainen 2015).

3.3.4.2 Confidence, error sensitivity and the predictive brain

The brain appears in several neurocomputational models as not only a probabilistic but predictive organ (Rao & Ballard 1999). The simplest predictive model represents the process of signal detection or learning in a hierarchical and linear way: faced with an input (stimulus), the brain makes a prediction; this is then compared with reality and produces an error signal (prediction error). This signal triggers an update of the initial prediction (learning). The larger the prediction error, the greater the correction can be. At the basis of any learning or signal processing, there would therefore be an initial prediction, belief or representation, and then an error signal if the prediction is incorrect (see: Rescorla & Wagner 1972; Sutton & Barto 1998). In the 1990s, the role of prediction (and the violation of a prediction) in learning is further supported by neuroscience data and models. For example, Schultz, Dayan and Montague (1997) identify groups of neurons that, in the case of reward and punishment learning, signal prediction errors (when the animal receives a reward, its dopaminergic system is activated; when it understands an association between stimulus and reward, this learning modifies the activation of the dopaminergic system, which then activates even before the reward arrives, thus signalling an expectation of reward). Following the arrival of a stimulus, the brain therefore makes a prediction that the reward *should* arrive soon, and this prediction is associated with some neural activation with dopamine discharge. If the reward does not arrive when it is expected, the dopaminergic system reacts with a depression of its activity below baseline. The system thus receives a prediction error signal in the form of a decrease in dopaminergic activity[144]. This is an example of the brain's ability to encode an error signal for use in decision-making in an implicit, sub-personal way.

Prediction/error/update is not specific to dopaminergic neurons or learning with reward, however, and its implementation varies across brain regions.

Integrative (predictive and Bayesian) models of brain function also highlight the importance of confidence in predicting and updating one's predictions - and thus in learning. When we are confronted with something new, we must indeed keep a subtle balance between what we have learned in the past (e.g. the observation of regularity) and new information, which can in principle call into question previous knowledge. Behavioural and neuroimaging studies indicate that the responses given during a learning task (conditional probabilities with perceptual stimuli) respond to an optimal learning algorithm that takes this subtle balance into account. Thus, low confidence in our *a priori* leads us more easily to update them in the presence of contrary evidence. Similarly, feedback associated with this low confidence has less influence on the revision of the prediction (Meyniel & Dehaene 2017; Heilbron & Meyniel 2019[145]).

3.3.5 Strenghts and limitations of metacognitive sensitivity: a matter of calibration

The discussion we have just presented shows us that our cognitive architecture is prepared to estimate the degree of confidence that we have in an information and that we can have in an opinion or decision[146]. Moreover, this estimation plays a key role in fundamental processes such as decision-making (in the broadest sense of the term) and learning (the updating of *a priori* and the expectations they produce). Confidence plays a particularly important role because we cannot systematically and temporally rely on corrective feedback of our actions: it allows us to track down and eventually correct our mistakes (Kepecs et al. 2008; Yeung & Summerfield 2012; Rouault, Dayan & Fleming 2018). Trust occurs not only at the personal level, but also at the sub-personal level, including in the context of low-level processes. We can therefore consider it as a structural element of our cognitive functioning.

However, this operation is not flawless. Studies on the correspondence between subjective confidence judgements and performance in adults (metacognitive sensitivity) show that metacognitive confidence judgements are often better than a random choice, but still not optimal, and often subject to error and bias (Metcalfe & Shimamura 1994; Dunlosky & Metcalfe 2009; Serra & Metcalfe 2009[147]). The dichotomy between metacognitive narrative and actual performance is an object of study in itself, as it demonstrates that decision and trust respond to different rules (at least in part), suggesting the existence of partially independent neural substrates (as highlighted in the discussion of trust and the brain) (Grimaldi, Lau & Basso 2015; Meyniel, Sigman & Mainen 2015).

Confidence calibration is defined as the relationship between the average of confidence judgements (on average, across the different tasks of a test) and that of performance[148]. Confidence judgments are considered to be poorly calibrated when they diverge from actual performance (high confidence in performance is accompanied by incorrect answers and low confidence is accompanied by correct answers). We also refer to the relative precision of judgements to indicate the ability of a judgement to predict the correct or incorrect nature of a test response (Serra & Metcalfe 2009). In addition, some authors invite us to distinguish between sensitivity and bias. Concerning the latter, during a test, a subject can make confidence judgements that are globally calibrated in relation to his performance, while tending to judge his performance in a rather optimistic way (Grimaldi, Lau & Basso 2015, see in particular: Fleming & Lau 2014).

3.3.6 Matacognitive biases, between over- and under-confidence

One of the most cited examples of metacognitive bias with negative calibration is central blindness: the subject solves visual tasks with greater than random success but does not feel capable of doing so. Other examples are represented by forms of implicit learning, where the subject begins to give correct answers before realizing that he has learned a new rule. In the task known as the *Iowa Gambling Task* (Bechara et al. 1997), subjects select packs of cards. The choice is initially random. Then, when the subject notices certain regularities in the decks, he or she chooses on the basis of the gains and losses that the different decks bring. We notice that subjects in the control group (without prefrontal lesions) start to choose correctly the less risky packs before they realize that they do so: for a certain period of time, they therefore choose correctly but they think that they choose randomly. Their explicit confidence is poorly calibrated, even though statistical learning still takes place (on the contrary, orbital-frontal patients do not learn or continue to choose disadvantageous packets even after they realize they are more likely to lose).
Optimistic biases are described at two levels (Sharot 2011; Sharot, Korn & Dolan, 2011; Sharot et al. 2012[149]): at the level of first-order cognition, they are manifested by the tendency to overestimate the probability of positive events and underestimate the probability of negative events; at the second-order level, they are manifested by an overconfidence in representations and predictions or decisions (believing that we are right, that our beliefs are true, our decisions are correct). The effects of overconfidence belong to this second - metacognitive - category, such as the "Dunning-Kruger" effect and the illusion of explanatory depth.

3.3.6.1 Over-confidence and the « Dunning-Kruger effect »

Several studies point to the existence of a bias that affects retrospective confidence judgements (Dunlosky & Metcalfe 2009).

The Dunning-Kruger effect is revealed when we measure the performance of participants in a variety of tasks, such as irony, logic, grammar... and ask them to rate their performance in relation to other participants: not only is there an overconfidence effect, but the overconfidence is also not uniform (Kruger & Dunning 1999). Studies show that we tend to become overconfident for more difficult tasks and in situations of underperformance: the better we perform, the more insightful we are about our own performance - or even underconfidence. We are therefore in the presence of a prediction error at the cognitive (first-order) and metacognitive (second-order) level, of the optimistic type. The Dunning-Kruger effect has been replicated several times (Dunning et al. 2003; Ehrlinger et al. 2008[150]) and has also been explored in the field of reasoning (Pennycook et al. 2017). The results confirm those obtained by Dunning and Kruger.

Dunning and Kruger explain that the recorded effect is due to the fact that the knowledge that is needed to realize that we lack knowledge is the same knowledge that is needed to complete the task. So when we lack the skills to do the task, we also lack the skills to realize that we lack the skills to do the task, because the skills are the same[151]. The explanation is consistent with other observations made in the field of expertise (for example, by Chi, Glaser & Rees 1982 on skills in physics). Studies on expertise, although not addressing the issue of self-assessment of competence, show that novices, compared to experts, lack certain metacognitive abilities related to the task, such as the ability to detect the actual difficulty of the task, and the time and effort required to learn from their success on the task[152].

This type of study and the explanatory hypotheses put forward highlight the importance of having "domain" knowledge in order to correctly assess the correctness of our opinions, decisions and performance. We will come back to this point in the discussion on strategies for CT education and skills transfer. We can anticipate by indicating that knowledge influences metacognitive abilities - including self-assessment, confidence in our decisions - which represents a difficulty for any educational approach to CT, because CT cannot be considered as disconnected from the content on which it is applied.

However, other characteristics of the task and its formulation influence performance. For example, and counter-intuitively, overconfidence may be inhibited in the case of particularly easy tasks (Lichtenstein, Fischoff & Phillips 1982). Thus, while difficult tests awaken overconfidence, retrospective judgements on easy tests activate an attitude of underconfidence. This effect is called the "hard-easy effect".

3.3.6.1.1 The Dunning-Kruger effect in reasoning tasks

The Dunning-Kruger effect has also been explored in the area of reasoning (Pennycock et al. 2017). Pennycock and colleagues used analytical reasoning tests such as the *Cognitive Reflection Test* and asked participants to rate their performance on the tests and to indicate on a scale their readiness for analytical reasoning. The results confirm those obtained by Dunning and Kruger. Pennycock and colleagues concluded that the ability to reason analytically also facilitates the metacognitive ability to monitor oneself during a reasoning task[153].

Nevertheless, it would seem that a certain degree of sensitivity to error - even in [154]the type of reasoning where we highlight recurring errors - is present at the sub-personal level and implicitly (Bago & De Neys 2019). De Neys, Rossi and Houdé (2013), De Neys, Moyens and Vansteenwegen (2010), De Neys, Vartanian and Goel (2008) have studied the confidence associated with negative performance in reasoning tests - in particular the confidence associated with performance in the classical bat and ball problem, integrated in the CRT developed by Frederick (2005). Participants giving the wrong (and expected) response to the bat and ball test also show lower confidence in their response to this test than to other control tests. De Neys and his colleagues have deduced from this that the detection of a conflict between correct and incorrect answers given is always present in the mind, even if it is implicit. De Neys, Lubin and Houdé (2014) have also shown that children who give the 'wrong answer' on a task called 'keeping numbers', however, have less confidence in their answer.

Pennycock suggests a possible solution for the discrepancy between the results of her experiment and those of De Neys' experiments: although the signal of conflict is present at the neurological level and can even be detected at the psychological level, it is not "strong" or effective enough to bend overconfidence. Indeed, according to Pennycock and colleagues, De Neys' tests show that confidence decreases in relation to tests such as bat and ball compared to others, but that confidence nevertheless remains, incorrectly, high (Pennycock et al. 2017[155]).

De Neys, Vartanian and Goel (2008) argue, however, that we are - even in relation to issues of trust and sensitivity to our biases - more 'intelligent' than we typically think, or that this is not reflected in traditional theories of bounded rationality[156] (we will return to this point in the next section).

3.3.6.2 The illusion of explanatory depth

Our cognitive functioning has been dominated, since early childhood, by epistemic intuitions - intuitions about the nature and laws of the physical phenomena that surround us, as well as [157] psychological and social laws (Rozenblit & Keil 2002). Babies divide the world (information) into objects by following certain intuitive and spontaneous "rules". They do not treat all objects on the same level but categorize them on the basis of certain properties: physical, biological, mental, social, mathematical objects, each category receiving its own treatment and being subject to a set of expectations: physical objects do not set themselves in motion, agents set physical objects in motion, objects do not disappear into the void... These expectations call for explanations. Children are very active in searching (rational exploration) and in detecting causes (identification of causal relationships based on simple temporal and spatial clues). Signs of these abilities and attitudes are present before six months of age and have been the subject of numerous studies in developmental psychology over the past thirty years. The term “*core knowledge*” is used to refer to earlier and more ingrained abilities and knowledge (although it would probably be better to speak of predispositions to learn); 'natural categories' to describe how concepts are formed; and 'naïve theories' to describe the fact that adults and children use intuitions to explain phenomena they observe or to answer questions they are asked about them (sometimes by inventing them, but not in a totally random way) (Spelke & Kinzler 2007).

However, we are unaware of the shortcomings of the explanations we provide. Our knowledge is limited, but we delude ourselves that we know more than we actually do, and in particular that we understand more than we actually do about the mechanisms underlying the functioning of objects of common use (Sloman & Fernbach 2018). Rozenblit and Keil (2002) study this phenomenon in adults. Their methodology is based on confidence judgements: participants evaluate on a scale of 1 to 7 their understanding of a phenomenon; then they describe the causal functioning of the object observed in detail and re-evaluate their understanding on the same scale; they answer a specific diagnostic question about the object and re-evaluate themselves; they read an explanatory text about the object and re-evaluate themselves, etc. (Sloman & Fernbach 2018). Finally, they answer a question about the usefulness of the explanatory text provided for understanding the phenomenon. The results indicate that subjects initially overestimate their understanding and that their confidence decreases after the effort to provide a detailed explanation. In the debriefing phase, the study participants declare their surprise at their poor knowledge. The study was then carried out by the same authors with other (but still homogeneous) populations and varying parameters. For example, in one test, participants were informed from the beginning that they would have to provide a detailed explanation of the functioning mechanism of the object observed: this reduced the difference between successive self-assessments, but did not make it disappear. The authors also tested other forms of knowledge and understanding with the same type of device: factual, procedural, narrative knowledge (the summary of a film). They found significant differences: overconfidence in facts, but less confidence in mechanisms, and good calibration for procedural and narrative. On the other hand, the same results are obtained for mechanisms and for explanations of natural phenomena. Causal explanations would therefore be particularly likely to generate overconfidence in relation to other knowledge contents. This would mean that overconfidence is not a general phenomenon but is relative to certain areas of knowledge[158].

Rozenblit and Keil (2002)[159] do not consider this illusion to be identical to the optimistic illusion that leads us to be overconfident in our knowledge and abilities - a phenomenon widely documented in quite diverse fields. The illusion of explanatory depth reflects a specific overconfidence in the understanding of causal mechanisms, and therefore in explanations. We have the illusion that we have more knowledge than we actually have, and that we understand the mechanisms underlying the functioning of objects of common use better than we do in reality. However, the same people who are overconfident about the accuracy of their causal explanations are not overconfident about other forms of knowledge, including factual or procedural knowledge. There are a number of reasons why this type of illusion may exist. We may feel that we understand the complex relationships between components of a mechanism because we understand how it works as a whole. Having even incomplete explanations may be of immediate use and may maintain the impression of usefulness in the long run, hence the term "explanatory depth illusion". Other reasons are also to be found in the feedback. For example, causal explanations about the working mechanisms of an object we manipulate get positive feedback from the moment we know how to turn the object on, but this feedback is misleading in relation to the gaping nature of our deep understanding of its working mechanisms. Moreover, in-depth explanations are rarely requested and therefore we lack feedback at this level (Rozenblit & Keil 2002).

3.3.7 Why might metacognitive sensitivity be miscalibrated?

There are several answers to the question of miscalibration of confidence, when it occurs.
First, optimistic biases and overconfidence may have adaptive value [160]. This would counterbalance the negative adaptive value of ignorance of real risks, over-optimism, and over-confidence that can lead to disastrous consequences. Sharot and Garrett (2016) note, for example, that the role of the positive value of feedback includes cases of motivated beliefs and ambiguous information. According to them, the tendency to give an optimistic interpretation to information is present in different species (birds, pigs) and is particularly evident when the situation is ambiguous and uncertain and in the absence of stress (Sharot & Garrett 2016).
Second, the prediction of success also depends on previous experience with similar tasks. In the absence of this experience, the child, and more generally the beginner, may be overconfident (Proust 2019[161]; Lockl & Schneider 2007).
Third, the fact that metacognitive feelings are the product of heuristics based on indirect indices of response validity can create pitfalls that affect reliability (Serra & Metcalfe 2009; Koriat 2012[162]). So what are the clues from which metacognitive feelings arise?

3.3.7.1 Heuristics and metacognitive illusions: the problem of the indirect character of cues

Serra and Metcalfe (2009) identify several heuristics at the origin of metacognitive illusions: familiarity with a certain material may influence the feeling of knowing it well; thespeed with which a memory is retrieved from memory may influence confidence in responses and other judgments related to the ability to remember and learn[163].
Koriat (2012) in particular proposes a model of confidence judgement based on memory availability. In this model, confidence in a response is correlated with cues such as the speed

with which the response is retrieved from memory or selected; it is also correlated with the accuracy of the response, but only indirectly, *via* information on speed. In other words, *ease of* retrieval is a frugal and indirect index for judging the accuracy of one's response[164]. The associations that are first activated by the task and the response depend on prior experiences and learning laws that make some associations are learned or reactivated more easily than others. Speed can therefore be a good clue, especially when the subject has a great deal of knowledge in the area of the task; but as an indirect clue, it can also lead metacognitive judgement into error.

Another clue used to judge, implicitly and frugally, the correctness of the answer is the coherence (*self-consistency*) between different representations present in the memory. We can see it in the following way: the confidence system does not try to measure directly the validity of the answer but the probability of giving the same answer in the future (*reliability*). Reliability is in turn an indication of the validity of the answer. The task activates a certain response and a series of representations. Each representation has a positive or negative sign in relation to the given answer. If the sum is positive, the confidence system detects a state of consistency and produces a feeling of confidence in the chosen answer. A miscalibration of confidence can occur by virtue of the indirect nature of the index[165]. Several aspects of the learning situation can also produce different forms of the illusion of knowledge[166].

3.3.8 Strategies for impriving the quality of metacognitive judgments

Even if they are natural, early and silent, metacognitive feelings such as sensitivity to error can be cultivated.

On the one hand, one strategy that seems to have a positive influence on the accuracy of metacognitive judgments and their calibration is to be aware of the biases that influence these same judgments[167]. This includes becoming aware of the heuristics and clues underlying confidence judgements, and understanding how these can produce illusions of understanding, learning, or giving the correct answer. The effect of testing with feedback would indeed improve performance but without necessarily producing a transfer to new material or new contexts. Explaining why we may fail would have this additional effect (Serra & Metcalfe 2009).

Butterfield and Metcalfe (2006) have uses "Trivial Pursuit" type questions to assess the ability to revisit an answer with a high degree of confidence. After answering the questions, participants estimate their confidence in each answer they gave. Their answers are corrected by feedback that tells them which answers were correct. Then the same topics are re-tested on the same questions some time later. Butterfield and Metcalfe (2006) have shown, first of all, that the presence of feedback makes it possible to change the answer, even in cases where the confidence in the answer had been very high. Second, that the questions on which participants were mistaken, and for which they felt particularly confident, are the ones that are corrected with the highest probability on the second test. This is the "hypercorrection" effect. This effect is counter-intuitive, as we would expect it to be more difficult to change the positions in which participants felt more confident.

According to the authors, the greater the error of confidence, the greater the surprise effect, and therefore the greater the attention paid to the error[168]. Butler, Karpicke and Roediger (2008) have tested the effect of feedback on responses that are correct but accompanied by a low level of confidence. They propose tests with multiple-choice responses. Subjects are forced to answer. This leads to situations where subjects may give correct answers by chance, but without feeling truly competent. These tests show that from a correct answer with a low level of confidence, the feedback improves first-order cognitive performance (knowing how to give correct answers on the second test) probably through an improvement in metacognitive ability[169]. In an experiment involving approximately 60 subjects, Carpenter and colleagues (2019) have compared the effect of feedback on the accuracy of response and the effect of feedback on the accuracy of metacognitive judgments about performance. The motivation for trying to directly influence metacognitive confidence without trying to influence first-order cognitive performance is two-fold: on the one hand, the authors report recent positive results in improving metacognitive performance obtained through interventions such as mediation, substance absorption or brain stimulation; on the other hand, they explain that neuroscience research has highlighted the existence of specific neural substrates for metacognitive performance, separate from the substrates involved in first-order cognitive performance (Carpenter et al. 2019[170]). We have cited above the positive results of interventions that propose to directly influence the calibration of confidence *via* explicit methods. In contrast, the intervention tested *via* the *Amazon Mechanical Turk* platform by Carpenter et al. (2019) is implicit, since it is based solely on the feedback proposal. The experiment includes a perceptual decision task and a memory task. In the perceptual task, participants judge the relative brightness of two abstract forms that are simultaneously proposed to them. The memory task involves the rapid visualization of a series of letters and the proposal of two alternative choices to recognize the initial image. Following their choice, participants make a retrospective judgement of their performance, declaring their confidence in the answer on a scale ranging from 1 (very low confidence) to 5 (very high confidence). After 27 trials, participants receive feedback on their performance (control group) and their metacognitive performance (experimental group) respectively. They continue in the same way for ten training sessions. The results indicate that, immediately after the first session, the participants in the experimental group see a decrease in their metacognitive bias, while the effectiveness of metacognitive sensitivity increases gradually but more slowly during the training sessions. Receiving first-rate performance *feedback* does not produce the same effect. Alternating perceptual and memory tasks makes it possible to assess whether the improvement of metacognitive abilities in one of the two domains is generalised in the other. The results indicate that a transfer takes place from perceptual training to performance in metamemory, but not for the metacognitive sensitivity component. The experiment conducted by Carpenter and colleagues (2019) does not allow us to know whether this type of gain is also transferred to real-life tasks, nor how long the benefit would last, but opens the way for further studies.

Depending on the nature of the (fairly simple) stimuli used in the study, we could hypothesize that the effects of training are stronger and more likely to be transferable if they are accompanied by instructions about the heuristics that influence metacognitive performance and if training is conducted on more realistic tasks.

The question of the nature of metacognition - context-dependent or context-independent - does not yet seem to be resolved (Rouault et al. 2018[171]).

3.3.9 Metacognitive confidence. Consequences for CT education

CT is - as we have defined it - the most accurate assessment of the quality of information for decision-making. However, this assessment does not simply involve an accumulation of clues as to whether information in our possession is reliable or not. It also results in an explicit state, feeling or judgment of confidence in that information. This information, in turn, influences our present and future decision-making and opinions.

Metacognitive sensitivity (sensitivity to error and implicit trust) appears indeed as a fundamental mechanism to take into account the value of information in order to make a decision, to regulate the ability to change one's opinion on the basis of new information and to update one's representations and predictions. Its functioning is active early in the child's life, but in its natural state, trust judgements - implicit and explicit - are based on indirect clues and heuristics. This dependence can lead to errors in judgment.

From these considerations emerge guidelines for the education of critical thinking. The main indication consists in equipping people to trust a certain piece of information or opinion. Metacognitive confidence can indeed be learned, and different methods are potentially useful. The literature on metacognition and trust thus provides not only a theoretical framework but also practical indications on how to improve these capacities, which are not without limits and are affected by illusions or bias like other capacities.

The literature reviewed allowed us to identify factors that positively influence the effective calibration of confidence:

- knowledge in the field concerned. Domain knowledge seems to be a condition for the proper exercise of confidence in the information in our possession, in our forecasts and decisions ;
- the presence of appropriate feedback. Feedback can be related to performance, but it can also be useful to provide feedback regarding the correct or incorrect calibration of trust and the presence of positive or negative trust bias, i.e. to inform the subject about the distance between his or her trust judgements and reality. To achieve this, confidence must be assessed and then compared with performance. It must therefore be part of the pedagogy of critical thinking. The fact of including metacognition in learning in an assumed way, and of equipping oneself with tools such as tests with feedback, has moreover shown its usefulness in other areas of learning. Its introduction on the side of critical thinking education can therefore only be consistent with indications given to educators elsewhere (see in particular: Proust 2009; Hattie, Biggs & Purdie 1996; Higgins et al. 2012);

- Confidence is associated with performance, but it is associated indirectly, *via* indices. One indication is to make these criteria and potential illusions of metacognitive judgement the object of explicit instruction, in order to identify and anticipate them. Explicit instruction seems to be a key element of the transfer, even if other strategies such as duly administered feedback are implemented;
- Since the use of indirect indices to judge the confidence that should be placed in an opinion or decision is subject to error, one strategy is to adopt more objective, external criteria. The literature on metacognition points out that external aids such as assessments and tests may be necessary to correct metacognitive illusions of knowledge and learning. This indication can be exported to metacognitive sensitivity in the following way: sometimes our internal metrics are insufficient to provide us with a correct estimate of the trustworthiness of information or representations that we hold. In this case, we want to adopt external scaffolds similar to those encountered in the first part: grids of criteria that allow us to objectify the quality of an information or another representation. Subjective confidence must then bend to these criteria, which the subject has recognized as superior to his natural indices by virtue of explicit instruction on the latter.

These tools are all the more necessary when we have to evaluate our own opinions, positions and decisions. Indeed, we have found that it is difficult to assess one's own competence and expertise. Yet this is a necessary condition for knowing whether we can trust ourselves when making judgements on any information. We can then assume that knowledge of cognitive and metacognitive biases represents one more tool, allowing us to better anticipate our chances of error, to identify situations that are more likely to put us in difficulty by creating obstacles to a fair judgment of confidence or reliability in information.

To conclude this discussion on metacognition, it is important to emphasize that the objective of equipping natural confidence mechanisms is to achieve a better calibrated confidence, a better sensitivity. This translates into the ability to declare oneself uncertain when the evidence at our disposal is of low quality, or lacking. For example, when the only evidence available is personal anecdotes, we should legitimately lower our confidence in the opinion or information concerned. This does not mean that we should refrain from forming an opinion, as this would be unrealistic, but that we should judge the opinion as weakly supported, and therefore as likely to be reversed, modified, in the light of new evidence. When the evidence is strong, and we correctly judge that we know how to value it, with our knowledge, then our confidence is legitimately high, and we can declare ourselves certain of our position.

We can represent this way of thinking about well-calibrated confidence as a continuum from "not at all confident" to "very confident". The CT exercise is about getting to the right place on the line, based on the information available to us, including assessing our own competence and expertise in evaluating that information.

Why is it important to make this explicit? Because CT is often described in terms of mistrust and associated with a defensive, protective, doubtful or vigilant image (such as the term "epistemic vigilance"). However, the discussion around confidence highlights the importance of being confident and wary wisely, and therefore of calibrating that confidence based on an assessment of supporting evidence, sources, plausibility and relevance of information, and our own competence. As thiss assessment becomes more sophisticated, more context-specific, using artificial tools that address natural capacities, the confidence also becomes better calibrated.

3.4 Seeking to understand the nature of errors

It is a fact that we make errors in the process of making decisions or forming opinions. We have seen that these errors can specifically affect the evaluation of information (errors of judgement of information content or sources) and the calibration of our confidence (overconfidence in a decision we have made, in a position, opinion, perception; or too low confidence, when our information is of good quality).

Moreover, our decisions, judgments, evaluations, perceptions are the product of silent inferences that take place below the threshold of consciousness, and that escape reflection. Often, attention is focused on how these implicit processes lead us astray. Thus, a vast literature in the psychology of decision and reasoning is dedicated to describing the biases that influence our choices, judgments, and opinions (see, for example, Pohl 2017; Kahneman 2011; Gilovich, Griffin & Kahneman 2002; Ariely 2008). This literature extends its success beyond its own domain and is often cited in the CT field to explain our most common errors

of perception, prediction, explanation, and more generally decision-making (see, for example, Battersby & Bailin 2013; Halpern 2013; Nisbett 2015; Stanovich & Stanovich 2010).
This literature belongs in particular to the so-called "heuristics & bias" program in social and reasoning psychology[172]. The idea behind this program is the following: under conditions of judgement in a situation of uncertainty, we resort to shortcut solutions that are limited in number, quick to implement and simple (Tversky & Kahneman 1974). These heuristics are generally useful but can sometimes be misleading. When a heuristic gives rise to a systematic error, we speak of bias (Kahneman & Tversky 1996). The wide variety of biases observed in the reasoning could be extended to a small number of functional heuristics (Kahneman et al. 1982; Gilovich, Griffin & Kahneman 2002).

3.4.1 The « heuristics & biases » program (H & B) and the nature of errors

The "H & B" program is rooted in the notion of *bounded rationality* introduced by Herbert Simon to describe the effects of the computational limits of our cognitive architecture on our choices and decisions. Its theoretical basis is thus constituted by the idea that our cognition is intrinsically limited and that efficient solutions can be developed to reduce computational costs and lead to satisfactory, not necessarily optimal, decisions[173]. Cognitive biases would be a manifestation of such solutions. The added value of studying cognitive biases for the psychologist consists in gaining access to solutions implemented structurally in our cognitive architecture. For the cognitive system that implements them, these solutions allow us to reach a certain level of efficiency, since they do not take into account the optimal solutions that we can find by taking the time and using more sophisticated reasoning algorithms adapted to the situation; they are also sources of error.
This program has been very positively received by a large part of the research community in the field of psychology of reasoning and enjoys wide success in various disciplines (medicine, economics, law, management, politics) and with the general public. (One need only think of the success of books such as *Thinking Fast and Slow* by Daniel Kahneman, one of the spokespersons of this program better known to the general public: Kahnemann 2011).

3.4.2 Criticism of the H & B programs

The programme has also been criticized, to varying degrees of severity. The main criticisms to the analysis of heuristics and biases come from approaches with a more ecological, evolutionary component, which have given rise to alternative programmes, such as that of adaptive rationality (Cosmides & Tooby 1994; Gigerenzer 1991, 1996, 2004; Gigerenzer, Hoffrage & Goldstein 2008; Haselton et al. 2009). This program is not limited to human cognition, but is also found in the field of ecology and animal cognition, where adaptive rationality of decisions is often studied in relation to foraging tasks (see Stephens 2008).

Among the most important critics of the H & B program is the psychologist Gerd Gigerenzer. He argues, among other things, that the way questions are asked by psychologists can create traps in which our cognition falls into by producing an error, whereas we would not be mistaken if the question were asked differently (Gigerenzer, Hoffrage & Goldstein 2008). Moreover, too much attention would have been paid to cases in which the cognitive system is defeated and led to make mistakes, or just fast but rationally unsatisfactory, and too little attention would have been paid to situations in which our cognition serves us well or even optimally[174]. In contrast to bias as the effects of "*quick and dirty*" solutions, Gigerenzer, Tood & The ABC Research group (1999) thus propose a set of heuristics that are rather "*fast and frugal"*: heuristics that make us more effective and efficient, that have an adaptive role, and that are part of a mental toolbox that has been selected during our evolution in response to specific tasks and problems to be solved. Decision-making involves choosing the most appropriate tool from the toolbox. An example is the recognition heuristic: let's imagine that we are asked which of two cities has the most inhabitants, we don't know the answer, but we have to base it on criteria. The one we would probably use would be the one of recognition (recognition heuristic): we recognize the name of one of the two cities so we answer that it is the biggest. The criterion in question serves us well in this case, because it is indeed related to the size of the city: larger cities are more often named on television, etc. However, like all indirect criteria, this one may not apply in some cases, or may be "activated" for the wrong reasons.

Adaptive evolutionary approaches criticize the "heuristic and biased" view of using only external norms to judge whether behaviour is rational or irrational. In the H & B approach, the norm is established on the basis of the desirability of a certain type of behaviour from an external, logical point of view. Whereas in the adaptive approach, we seek to establish whether the norm really corresponds to the objectives that the cognitive system responds to in order to increase the individual's selective value under given conditions. The evolutionary approach thus takes into account the ecological conditions in which behaviour emerges[175]. For example, while truth-seeking behavior responds to culturally important norms, there is no assurance that the solutions selected in evolution meet only such norms. This selection has taken place under multiple constraints - including, but not limited to, truth-seeking - and in complex environments, which we often neglect in a classical H & B approach.

3.4.3 The "evo-eco" approach: putting our intuitions and instincts in a positive light

A fundamental concept of the evolutionary and ecological ("evo-eco") approach is that of a *match* (and *mismatch)* between the ready-made solutions (heuristics) that we implement and the environment in which we act in the present and the past. The solutions available in our cognitive toolbox are the fruit of our evolution; their structure and functioning are therefore determined by the conditions of our past where their evolution took place (ecological constraints). A certain heuristic is neither rational nor irrational in itself: it is adapted to a certain context and gives an optimal response in that context. The apparent sub-optimality of a strategy can then be explained in different ways: it is possible that it is a physiological limit and that the tool could or could have become more efficient, but other causes can be envisaged. For example, the tool was optimal in the past context in which it was selected; it is the best compromise when we consider all the constraints acting on it, not just the one that interests us immediately. If we want to properly judge a heuristic (or any biological trait), we must therefore imagine the original context of its use[176].

3.4.3.1 An example of the evo-eco approach to decisions and behavior

To better understand the shift in perspective that an evolutionary approach to evaluating a choice of strategy (decision-making) implies, we will move away for a few moments from human concerns related to decision-making and rationality to behavioural studies of other species. This comparative approach allows us to hypothesize about seemingly "irrational" human behaviour.

In behavioural ecology, so-called "*Optimality Foraging Theory*" models are used to determine the constraints that act on an animal when making food-seeking choices. This approach makes it possible to better characterize the adaptive value of behaviour. For example, studies conducted on crows (Richardson & Verbeek 1986) show that crows leave part of their prey (clams) intact after locating them. Why is this potential food abandoned? Is it a loss of energy due to irrational behaviour? Crows accept 100% of prey of a certain size and the proportion decreases below that size. By considering the caloric benefits of prey of different sizes and the costs of searching, digging, opening the prey and feeding, the researchers constructed a mathematical model based on the assumption that birds would maximise the calories they take in: an optimal diet. The observations match the model's predictions. The researchers also established a link between the selective value and the gain in energy gained per day. This therefore validates the approach: assuming that the behaviour was adaptive, and without too much *preconceived notions about* evolutionary constraints, they revealed the ones that actually affected the animal's choice and determined that its behaviour was really optimal. At other times, the model's predictions deviated from observed reality. If factors other than food intake influence prey selection, then the caloric maximization model is rejected after the test. It is then possible to add other parameters, such as the presence of predators. Some constraints may also apply to only part of the population.

Ecological models of optimality therefore do not seek to detect irrational behaviour in relation to pre-established economic norms. The assumption is that natural selection has optimized behaviours on the basis of their selective value. The strategy therefore consists of establishing the constraints that have shaped the behaviours observed today (Stephens 2008). Concluding that a behaviour is "irrational" in a given experimental setting therefore implies first understanding the distal causes of the behaviour in question.

This approach can be imported into the problem of human reasoning and decision-making. Gerd Gigerenzer (1991) assumes, for example, that failure to assess the prevalence of a disease (neglecting the baseline rate - Casscells, Schoenberg & Grayboys 1978) is related to the fact that participants must assess the probability of a single event. Some of our reasoning mechanisms include aspects of probability calculus, but they are designed to work with frequency information to produce frequencies. Reformulating the Casscells problem in terms of frequencies and imposing a frequency response significantly increases performance (76% vs. 18%). Hence the need to look at the reasoning modalities from an adaptive perspective by asking: what problems have humans had to face during their evolutionary history? Why, for example, have they not evolved to deal with these probabilities of individual cases? It may be because frequency-based estimates are useful for decision-making.

On the basis of the considerations expressed above, other trends that are considered to lead to irrational decision-making - risk aversion, for example - can be reinterpreted as very effective strategies depending on whether we consider one optimality rather than another (e.g. minimizing the variance of earnings - unless strong earnings are required - rather than simply maximizing earnings) (Cosmides 1989; Cosmides & Tooby 1996; Gigerenzer 1991; Gigerenzer & Hug 1992; Oaksford and Chatter 1994; Rode et al. 1999).

Let's take the following example: by offering participants two boxes, one containing 50 black and 50 white balls, the other containing 100 black and white balls of unknown distribution, they choose the first box twice consecutively to find a black ball (thus assuming that the probability of finding a white ball is lower) and then a white ball (Rode et al. 1999). This is interpreted in terms of risk avoidance (always take the box that minimizes the risk), against all natural reasoning (it is not possible that box A is both favorable for finding a black ball and a white ball since the assumptions are mutually exclusive). According to the authors, it is more likely using this strategy to win twice consecutively (1/4) but less likely not to lose twice. A need equal to one win should push us towards the "risky" strategy while a need equal to two wins should push us towards the "risk avoidance" strategy. This type of experimental approach shows that the participants possess strategies that are rational, not from a mathematical or logical point of view, but from an ecological point of view: well adapted for the resolution of adaptive problems, which our ancestors faced during our evolutionary history (minimizing variance except in the case of high need when searching for fruit in uncertain situations). What we consider to be a bias leading to irrational behaviour can in fact be completely reinterpreted as a "*better than rational*" strategy.

3.4.3.2 An example of the evo-eco approach to perceptual illusions

The notion of error has been criticized even in areas other than reasoning. Cognitive illusions, or bias, are indeed often defined by analogy with perceptual illusions as systematic (deviations from reality) errors (both predictable and robust at the inter and intrapersonal level), involuntary and difficult or even impossible to control, surprising. Cognitive illusions are considered to be more problematic to define in relation to the perceptive, especially in terms of the notion of error. Indeed, it would be more difficult to establish the presence of an error because the "norm" against which the error is defined is less objective. Cognitive illusions would also be less "inflexible" than perceptive ones, at least some of them because learning and manipulation of the situation could make them disappear, which is not the case with perceptive illusions[177]. In fact, the idea of error has already been debated in the case of perception. As in the case of reasoning, the ecological current of the psychology of perception considers the notion of error and illusion to be misleading. Again, our perceptual systems are the result of natural selection that has led to functional systems in given ecological contexts. As with reasoning, psychologists impose standards on our functioning that are not those that guaranteed the survival of our ancestors. For example, perceiving "correctly" the shades of colour of a berry in the middle of the foliage is less functional than perceiving the berry as having the same colour regardless of the time of day - (even though changes in the medium (the air) cause the colours to be physically different in the morning and evening). Our haptic system responds to physical quantities that are not necessarily those captured by modern terms of weight or mass, but rather "mixed" qualities that take into account both the mass and volume of an object. What a "cognitivist" psychologist diagnoses as a perceptual error - an illusion of color or an illusion of weight - is actually, for a psychologist in the ecological tradition, a sign of adaptation that has worked well: the perceptual system responds to objectives and physical qualities that are functional to the survival of the organism.(For a comparison of "cognitivist" and "ecological" positions to perceptual illusions, see: Pasquinelli 2006 (doctoral thesis) and Pasquinelli 2012).

3.4.3.3 Where do our errors come from?

The outcome of the "evo-eco" approach is as follows: individuals reason well when the circumstances are favourable (in the sense that they are consistent with the conditions that allowed a selective advantage to the module in question). This adaptive rationality (as opposed to a certain vision of the "irrationality" of our reasoning) should motivate us to understand the origin of our errors rather than simply assuming our permanent irrationality[178]. In what follows, therefore, we will look in more detail at how evolutionary theory explains the existence of biases, and distinguishes several types of biases according to their origin in adaptive terms[179].

Indeed, even within the framework of an "evo-eco-approach", the concept of bias remains legitimate. It is used to indicate deviations from an expected standard - which corresponds to objective reality (Haselton, Nettle & Murray 2015[180]; see also, in another area, Lea & Ryan 2015). Like the concept of illusion, the concept of bias is also useful at least at two levels: it allows us to identify conditions (context, stimuli) where our cognitive functioning systematically takes a "surprising" direction; it allows us to better understand certain structural aspects of our functioning (what kind of stimuli or classes of stimuli we respond to in certain contexts)[181] (see: Pasquinelli 2006 (doctoral thesis) and Pasquinelli 2012 for a discussion on the usefulness of the concept of illusion, despite the criticisms of the ecological current).

Haselton, Nettle and Murray (2015) have identified three possible reasons why biases may emerge as "design flaws":

- the solution works well in most circumstances, but it is not optimal, just satisfactory because of our limitations ;
- the error is actually on the side of the experimental psychologist, who measures the response to a certain type of task, whereas the function being measured responds in kind to other goals; and for these goals, it is precise and effective ;
- the error has a very low cost in kind compared to the cost of an error-free response, so it can pass the natural selection filter[182].

Type 1 error is the most discussed in the bias literature. Whether it is a "quick and dirty" or "fast and frugal" solution, a heuristic with limitations. Moreover, this kind of quick and cheap solution exists precisely because we have processing limitations[183].

Type 2 error is the error of forgetting to consider the fact that ecological optimality may have a different meaning than that established by the experimenter (ecological optimality, context-related vs. external norm, logical). This category also includes situations in which old evolutionary responses are unintentionally triggered by recent (evolutionarily salient) stimuli that resemble the natural stimuli that activate the response - possibly the stimulus may present itself as a superstimulus: a stimulus more powerful than the natural one[184].

Type 3 error is based on the assumption that errors can result in false positives or false negatives: what matters is the cost of these errors. Instead, a well-adapted system will allow errors that do not significantly affect *fitness*, while errors that are disastrous for fitness will be corrected (Haselton, Nettle & Murray 2015; Haselton & Nettle 2006[185]).

3.4.4 External and internal standards: back to CT

Treating the notions of bias and heuristics from an evolutionary perspective allows us to consider rationality from an adaptive rather than a normative point of view. Does this mean that we should therefore abandon the idea of external norms or standards against which to measure human performance?

This would make it very difficult to define goals for CT education, and correlated evaluation criteria. Education (at least formal education) has indeed a normative purpose, linked to the cultural developments of our species. It is a form of engineering, which aims at modifying natural capacities, equipping and developing them in directions consistent with culture.

The naturalization of bias, the notion of adaptive rationality and, more generally, the evo-eco-approach are not inconsistent with establishing cultural norms of desirability or giving greater value to the criterion of respect for truth and adherence to facts. These norms constitute objectives to be achieved and allow for an assessment of whether the objectives are being met (Haselton et al 2009[186]).

However, if the criteria are established "externally", what is the point of mobilizing an adaptive and evolutionary approach? Could we not limit ourselves to educating CT on the basis of a number of established standards, normative objectives, and ignore the aspects we have presented so far?

As we have seen, unlike the H & B approach, the evo-eco- or adaptive rationality approach aims to explain why it is so difficult to comply with some of these culturally established norms and standards (gap between external and internal norms, or type 2 errors). The evolutionary approach further recognizes that intuitive solutions that meet internal norms may have limitations and lead to error (Type 1). Errors that are more likely to have passed the screening filter are those that do not involve a vital issue but are borderline super-prudential errors (Type 3).

Although the motivations may appear to be of purely theoretical interest, adopting an "evo-eco-approach" has direct consequences for the education of CT and also for our pursuit of its naturalization.

The theory of evolution provides a framework for analysing the biases and errors identified by the psychology of reasoning, in particular *via* experimental devices used to explain behaviour in ecological situations. By better understanding the causes of the choices made by the learner (in particular by distinguishing between proximal and distal causes: the influence of the context, the way the problem is posed, the implicit goal pursued by the learner), we can better evaluate the limits of his reasoning and envisage margins for progress. In this way, we avoid certain pitfalls that would lead to inappropriate advice or educational paths.

The "evo-eco" approach shows in particular that we are not in a fight against the biases and errors of our cognition. Rather, it is a question of trying to identify the cognitive capacities that naturally enable us to evaluate the epistemic quality of information and to make decisions that conform to reality (or rather optimal decisions, according to the definition given of optimality). These capacities are not without limits. The adaptive approach even allows us to anticipate the contexts that are likely to challenge them: those that move away from our natural evolution, such as new information contexts, or the needs of our culture. We can then seek to improve each of the capacities identified with the help of tools that we summon when situations require it.

3.4.4.1 "Don't throw the baby with the bathwater": the value of an H & B cognitive approach to CT

However, it would be excessive to oppose too clearly the two approaches "quick and dirty" and "fast and frugal". Evolutionary criticism of H & B approaches rather highlights the need to question the origin of errors and biases that affect decision-making and the formation of opinions and beliefs, in order to target realistic improvement strategies rather than simply admitting these biases and seeking to counter them.

This type of criticism does not eliminate the fact that, in certain contexts (type of question, type of example), we make "mistakes" (see the responses of Gilovich & Griffin 2002 and

Kahneman & Tversky 1996). Gigerenzer himself admits that real cognitive illusions also exist and that the type of "treatment" we can apply to the more artificial ones reduces them without necessarily making them disappear: something still needs to be explained, i.e. how our cognitive functioning is induced to give an answer that it would not have given in other circumstances (Gigerenzer, Hoffrage & Goldstein 2008). The opposition between the two approaches "quick and dirty" and "fast and frugal" thus seems excessive in light of the facts, as we have the impression that what is most criticized is the fact that the focus of attention is on one aspect rather than the other of the same phenomena.

3.4.4.2 What can we do with biases and errors in a theory of CT?

Two processes are often indicated in the literature on reasoning as fundamental for making a correct, unbiased decision:

- slow, costly thinking (as opposed to automatic, fast, economical thinking processes) ;
- inhibition, or control, over responses based on more automatic thought processes.

An extensive literature in the psychology of reasoning describes cognitive functioning as based on two types of processes: Type 1 (T1) processes, which are fast and effortless, and Type 2 (T2) processes, which are slow and require an effort in terms of otherwise limited cognitive resources. Speed and effort are the characteristics associated with both types of processes, but often T1 and T2 are also opposed on the basis of their relationship to intentionality, controllability, access to consciousness and efficiency (Pennycook et al. 2018).
In the 1970s and 1980s, *dual process theories (*DP) aligned long lists of characteristics always associated with each other (if a process is voluntary, it is also conscious, and so on). Current PD theories have abandoned these long lists of characteristics. Evans and Stanovich, for example, who belong to the most influential representatives of PD theories, separate the defining characteristics of T1 and T2 from the characteristics that can be potentially associated with them (Evans & Stanovich 2013). They identify as defining characteristics of both types of processes: whether or not they require working memory and whether or not they are autonomous - whether or not they require attentional control. The authors also recognize that the lists of characteristics in traditional PD theories are often inconsistent: for example, Type 1 processes are not necessarily evolutionarily older than others; Type 1 processes may lead to effective and efficient responses, while Type 2 processes may lead to biases and

errors. How do Type 1 and Type 2 processes relate to each other? According to one of the dual approaches, the "default-interventionist" approach, when we are confronted with a new problem, intuitive responses are activated automatically and effortlessly. If we are not prepared for this type of situation, then the responses may be inappropriate. We must therefore intervene in a reflexive way to go beyond the intuitive response. However, we are often lazy: we choose clues that are easy to identify and recognize and do not look for the more precise characteristics of the situation. This leads us to choose the wrong answer (Evans & Stanovich, 2013).

Several criticisms can be levelled at dualist models, particularly those of the default-interventionist type (Mercier & Sperber 2011; Pennycook et al. 2018; Bago & De Neys 2019). Recent experimental data cast doubt on the two hypotheses of bias blindness and the purely corrective nature of "algorithmic" processes (T2). Several studies have shown, for example, that even when they fall into "bias" during classical logic tasks, subjects can detect the existence of a conflict, and thus be sensitive to error (Bago & De Neys 2019; De Neys, Lubin & Houdé 2014[187]). Indeed, confidence (measured by subjective and explicit verbal scales, proposed after the answer has been given) is lower in the case of tasks with conflict than in tasks without conflict, even for subjects giving the wrong answer. Other results show that a majority of the subjects who give the correct answer after reflection (slow condition) have already given the correct answer in the fast condition. In the case of these participants, the initial intuition is (and remains) correct. These data invite to rethink the relationship between T1 and T2: T1 could give intuitively correct answers even from a logical point of view; T2 would rather allow to justify *a posteriori* and to communicate one's intuition (Bago & De Neys, 2019).

As a result, rather hybrid rather than dual models of thinking have recently emerged (Bago & De Neys 2017, 2019, but also: Ball, Thompson & Stupple 2017; Białek & De Neys 2017; Pennycook 2017; Pennycook, Fugelsang, & Koehler 2015; Newman & Thompson 2017; Trippas & Handley 2018). In these models, a "rational" response may emerge at the outset, at the level of Type 1 processes. The choice between competing intuitive responses depends on the strength with which they are respectively activated (Bago & De Neys 2019). Since both responses are simultaneously “present in mind”, a conflict is detected between the two. The subject experiences a feeling of confidence in the given response, the stronger the less activated the other response is. This hypothesis helps to account for inter-individual differences: different subjects may have more or less strong logical intuitions, depending on how easily their logical response is activated. This hypothesis is also in line with the argumentative theory of reasoning proposed by Mercier and Sperber (2011, 2016) and Trouche, Sander and Mercier (2014). The role of the T2 operations would be to allow the user to find a justification for his or her choice of response, and to communicate it to others, possibly in order to convince them.

Another question that arises in the context of dual (including hybrid) models of thinking is how we move from one type of functioning to another (actually from T1 to T2 functioning), and what determines the winning strategy (the one that will determine the decision).

A classic answer to the question of bringing the more automatic and reflexive processes of thought under control, in favour of the more reflexive ones, brings the mechanisms of inhibition into play. Rather than speaking of two systems, Stanovich then proposes a model in three typologies of processes: fast T1 processes; T2 processes are algorithmic and can substitute for T1 processes; T3 processes, that are reflexive, and which control the implementation of T2 processes and inhibit T1 processes. This third type of process is more a disposition than an ability: it is a disposition to seek more information, to take different points of view, etc., and thus to allow algorithmic cognitive abilities to take place (Evans & Stanovich, 2013).

3.4.4.3 Inhibit?

The concept of inhibition makes it possible to think about the relationship between the different strategies coexisting in the organism for solving a problem, but it is also problematic (Aron 2007).

In its primary meaning, inhibition is a neurophysiological process involved in the central control of motor reflexes and the property of neurons that communicate [188]*via* particular classes of neurotransmitters, the inhibitory neurotransmitters (e.g., GABA) whose action inhibits that of the target neuron. The effect of these cellular processes is therefore that of a modulation of neural activity. The concept of inhibition is also present in psychology, with meanings that do not necessarily coincide with those of neurophysiology (Aron 2007). In particular, inhibition is treated as an executive-type function, sometimes called "cognitive control". Executive functions are defined as higher-order functions, responsible for controlling the implementation of lower-level functions. There is no single classification of executive functions, which may include several of the following functions: executive control, working memory, attentional control, cognitive flexibility, planning ability. One influential classification, however, reduces them to three: inhibitory control, working memory, cognitive flexibility (see Diamond & Ling 2016).

Executive control may be demonstrated experimentally by specific tasks, including negative priming tasks. Let's imagine that a subject has to react to an O1 object - a rabbit - and ignore an O2 object also present - a cat. In a second trial, immediately following the first, the subject must respond to a new object O3 - a dog - or to the previously blocked object O2 - a cat. In the latter case (O2), but not in the case of the new object (O3), the response is slower. This characteristic of the response is interpreted as a residual effect of a prior mechanism of inhibition of the O2 response (Neill, Valdes & Terry 1995, Tipper 2001). Here, executive control acts on attention, especially selective attention. More generally, the role of inhibitory or cognitive control would be to enable decision-making (including decisions at the sub-personal, perceptual or motor level for example) without being completely dominated by external stimuli that arrive continuously and can automatically trigger response actions (Tipper 2001). Patients with damage to the prefrontal cortex, for example, exhibit behaviour that can be interpreted as a lack of inhibitory control, as they tend to grasp any object that is placed in front of them (Lhermitte 1983). An important seat of executive control would therefore be in the prefrontal cortex (PFC). Neuroimaging studies have highlighted the specific role of the lower frontal cortex, especially the lower gyrus (IFG), in logic type reasoning and in its perturbation (Aron, Robbins & Poldrack 2014). In parallel with studies of reasoning and its neural underpinnings, Adele Diamond and other developmental psychologists have used the concept of inhibition - lack of inhibition - to explain certain child behaviours such as responses to the Piagetian test A - non-B (the child repeatedly searches for and finds a hidden object in place A. When the object is moved in front of his or her eyes to place B, the child continues to search in place A) (Diamond 1991).

However, it should be noted that the concept of inhibition is not self-evident in the psychological literature and is the subject of several criticisms.

Firstly, as presented, particularly in the context of dual theories of thought, it implies the existence of some kind of superior entity within the brain functioning that would trigger inhibition in order to bring the automatic response under control and give free rein to the algorithmic one. The three-system model proposed by Stanovich only postpones the problem of triggering inhibition.

Second, the emergence of logically correct responses can also be explained by mechanisms other than inhibition: for example, in the hybrid theories discussed above, by mechanisms to reinforce desirable responses or relevant information. This reinforcement would allow the correct or desirable responses to pass the decision threshold and thus to be chosen, even if others - normally more automatic - exist (Bago & De Neys 2019).

Third, even the effects of PFC lesions can be explained by mechanisms other than inhibition, for example, by postulating that these lesions involve disruptions in the functions that allow the task to be performed, rather than disruptions in an inhibitory system (Aron 2007). Some authors therefore consider the concept of inhibition to be rather a shortcut used to describe situations where two or more possible responses are in conflict. This conflict is then detected and a choice is made, with the involvement of working memory; the mechanisms allowing this choice and the development of a good performance are not necessarily related to the inhibition of one of the responses or stimuli involved. The use of the term "inhibition" in all these situations could lead to confusion because, for example, it conceals the existence of *top-down* modulating mechanisms that act more like amplifiers than signal inhibitors. Inhibition should therefore be considered as one of the mechanisms involved in this type of modulation and response control, and the concept of inhibition should only be mentioned in appropriate circumstances. Among these, motor control, but also control over emotions, attention and memory seem to bear the signs of active inhibitory processes, which block responses already initiated at the central level (Aron 2007). Neuroimaging seems to be a particularly suitable way of distinguishing different mechanisms potentially involved in the modulation of responses observed at the behavioural level (Aron 2007).

Fourth, efforts to "train" a possible general inhibition capacity through "classical" activities - such as the Stroop task or the *Go-no-Go* task - do not yield unambiguous results, especially when we are interested in transferring training to more or less distant tasks (Enge et al. 2014[189]; see also: Simons et al. 2016). Thus, in a meta-analysis of 19 studies concerning inhibitory control (*via* training with classical tasks such as *Go-no-Go)*, Allom, Mullan and Hagger (2016) found a small positive effect with transfer on health-related behaviours, but on transfer tests proposed immediately after training[190]. Diamond and Ling (2016) dedicated a literature review to the different interventions concerning executive functions, and found regularities with respect to the interventions that proved to be effective. The first consideration concerns the fact that executive functions can be improved, and that effective interventions are as much cognitive interventions as interventions that involve physical activity (martial arts for example). Other considerations relate to the type of effects and how to carry out effective activities. The transfer is narrow: training an executive function such as working memory does not provide benefits in terms of inhibitory control or cognitive flexibility. Those who start from low performance gain more. Time also plays an important role: prolonged practice is always more effective than short-term practice; moreover, when practice stops, the effects may disappear. Effort also plays a role: continuing to practice once the task has become easy has no input, as the task must remain a challenge to have an effect on executive functions (as with gaining expertise in general). An activity intended to train executive functions that has proven effective in a certain context is not necessarily always effective, since its effectiveness depends on the way it is carried out and practised: the same intervention may be effective in one context but not in another. Thoughtless physical exercise is not effective. On the other hand, physical exercise accompanied by reflection, planning and problem-solving shows positive effects. As for cognitive training, the best known is working memory (CogMed). It has shown positive effects (but not remote transfer and on real life tasks) but only in the presence of a monitor that accompanies the computer training, making it reflexive. This leads to further doubts about the effectiveness of trainings whose characteristic is to be implicit, not reflexive[191].

Houdé and Borst (2015) study - using neuroimaging methods - the role of inhibitory processes in logical reasoning (*if-then* rule in perceptual tasks) and show activations compatible with the hypothesis of inhibition. Houdé and Borst (2015) also show that instruction that involves negative feedback on the wrong answer is more effective - promotes the appearance of the right answer - than simple positive instruction that aims to reinforce the right answer. The authors call this negative feedback "inhibition"[192]. However, the feedback is explicit, and the authors therefore consider that explicit metacognition (which promotes the development of the inhibition process) plays a role in the success of instruction.

In conclusion, we can legitimately consider the concept of inhibition as the only process, or function, involved in the modulation of behaviour, especially in the presence of choices to be made between responses, and as a single concept. We can also consider the effectiveness of implicit inhibition training as a method for reducing errors in reasoning at different levels and in different contexts, unless it takes the form of explicit metacognitive teaching. Nevertheless, in light of the present knowledge, the narrow concept of inhibition and studies on specific inhibitory control mechanisms remain of interest for an operational theory of CT, especially when coupled with an explicit metacognitive approach.

3.4.4.4 Become aware of the most common biases and of the circumstances that can put us at risk

We have established that biases are not in themselves errors that must be systematically eliminated. Inhibition training that would make students globally less likely to follow their intuitions or automatisms would therefore be undesirable. We have just seen that this type of training, thought of as an implicit form of inhibitory control gymnastics, is not effective either. Two questions remain open. The first concerns the pragmatic utility of an awareness (explicit, metacognitive) of the limits of our own cognitive functioning, where errors are more frequent or damaging. The second analyses the effects on the ability to correctly assess the epistemic quality of the information at our disposal, and to calibrate our confidence accordingly.

Let us take the concrete case of a common reaction, which consists in interpreting - perhaps too quickly - a correlation relationship as implying a causal connection. The correct assessment of the quality of available information requires inquiring how the causal relationship is established, by methods that may or may not reduce uncertainty. However, realizing that we can easily be induced to use the correlation index alone as a causal factor could help us understand why appropriate methods are needed, how they reduce the chances of error in a causal judgement, and thus motivate us to use them or to take them into account in our judgements of information quality. Is this the case?

The literature on informed decision-making is, again, not very optimistic about the effects of entrainment (Marewedge et al. 2015[193]). It tends to show that transfer from one domain to another is a real challenge (we will return to this issue when we discuss the educability of CE). What about so-called "*debiasing*" (Fischoff 1982)? These most often consist of exercises that help us understand that a certain type of information tends to be disregarded when we make judgements. The most famous case of success is the debiasing activities proposed by Richard Nisbett and his colleagues in the case of teaching statistics, probability and cognitive bias to medical and psychological students (Nisbett et al. 1987). These interventions have been evaluated using tests, at time intervals and outside laboratory conditions, and have shown transfer effects on tasks that require the application of statistical and logical tools learned in class. There are also successful cases with a single intervention (video illustrating bias and simulation activity). Success is measured immediately after the intervention and at a distance but *via* close transfer tasks (Marewedge et al. 2015). Aczel et al. (2015) have implemented a training in several phases: the learner is first confronted with his errors *through* a specific test and *feedback*. Then he learns to recognise similarities between different situations and to extract a common principle: bias. He is then pushed to look for the equivalent in his autobiographical memoirs and instructed about the nature of the bias. At this point, he also learns strategies for responding to it and must look into the future to find situations in which he will reuse this learning. The post-test takes place one month after the training and uses the same questions on which the subjects were initially tested to discover their biases. Aczel and colleagues (2015) also analyse different categories of de-biasing techniques and highlight the difficulties generally associated with transfer[194]. Teaching about the existence of bias that affects decision-making, especially when it remains at an abstract level, does not necessarily allow the learner to recognise the concrete situations in which the bias occurs and to know how to react. Learning with examples and in specific application areas is linked to learning in these contexts, especially over time. The surface similarity between learning situations and application situations facilitates transfer, but this condition sets aside many situations or contents that are not considered. Moreover, learning the nature of bias without developing strategies to overcome it (for example, strategies linked to the application of statistics that are more sophisticated than those used spontaneously) does not allow for an appropriate response, even if the bias is recognised. We will discuss possible solutions to facilitate the transfer in the section dedicated to the educability of CT, but the objections raised here suggest that de-biasing is only effective under very specific conditions.

3.4.5 Heuristics, recurring errors and biases: guidelines for CT education

CT - as we have defined it - is the process of assessing the quality of information as accurately as possible and forming a judgment - implicit or explicit - of confidence calibrated against that information for the purpose of making a decision.

This definition is deliberately restricted. However, it leaves room for the search for specific cognitive components or *building blocks*. We have previously identified two cognitive components of CT: the processes and mechanisms of epistemic vigilance and those of naïve epistemology and metacognitive sensitivity. The latter consists in particular in assessing confidence in an information or decision on the basis of available evidence. This constant assessment can generate feelings of confidence at the personal level, and inform judgement. However, it can also translate into a characteristic of neural activations only in relation to certain stimuli, and thus only act on a sub-personal, functional level. The discussion concerning the notion of error and bias shows once again the importance - in assessing the quality of information - of taking into consideration one's cognitive functioning, and not only one's own.

From our discussion of heuristics and bias emerges in fact the consideration that in human cognition operate universal attractors that make certain information more attractive, certain ideas more convincing, or more understandable, attractive, salient, memorable - and this independently of their truth value. **Becoming aware of these attractors, and knowing more about cognitive functioning (not just one's own, but the functioning of cognition in general) would make it possible to better anticipate the chances of making a judgment error with regard to the information available and its quality.**

Does this mean that CT consists in learning to tame all our biases, to resist, to inhibit our spontaneous and rapid responses, i.e. our intuitions? As mentioned above, what we commonly call "biases", with negative intent, lead to errors only in certain contexts and circumstances. A general inhibition of biases or of the functioning of the so-called "type 1 of thinking" is therefore not desirable *per se*.

However, **if taken in a broad sense, the concept of inhibition can also include strategies and knowledge that do not directly block an undesirable response, but increase the chances of a desirable response. These strategies and knowledge must first and foremost be present in our minds, so they must have been learned if we are to hope to use them at some point in time. Moreover, their access must be easy, so they must be automated or almost automated in their use, in order to be used at the lowest possible cost**. We have seen the example of knowledge in statistics, which are tools to counteract recurring errors in our spontaneous appreciation of the role of chance, of the risk linked to events.

The question remains as to the importance of knowing how to recognize different types of bias in order to anticipate them and thus reduce their possible negative effects on the evaluation of the quality of the available information and the calibration of confidence for decision-making. While the literature gives us a sceptical view of the chances (and desirability) of generally debilitating human cognition, it does leave room for **educational interventions to become aware of biases and to use metacognition to better identify situations where the risk of making assessment errors is greater.**

An example will serve to clarify this point. A vast literature is dedicated to confirmation bias: the tendency to look for information that is consistent with our opinions and to evaluate it more begninninarily than we would with information of the opposite sign, plus a resistance to change our opinions to embrace those of others. We have described uncertainty - and therefore lack of confidence - as a driver for changing positions and overcoming confirmation bias. However, knowing the existence and strength of this bias also allows us to force our hand, voluntarily. For example, by creating communities of peers who evaluate and criticize our arguments for us. This is the case with the creation of consensus groups, scientific committees of peers, but also cooperative working groups. Knowing the underlying mechanisms of confirmation bias and other biases of a social nature allows us to anticipate that the groups that will work best in terms of overcoming confirmation bias are heterogeneous groups where contrary opinions exist. This prediction is also confirmed by some studies on group reasoning (e.g. Mercier 2016), although this knowledge needs to be linked to others regarding prestige effects, compliance, etc.

From these considerations emerge practical guidelines for CT education. The principal one is to **equip the capacity to assess the quality of information and the capacity to calibrate confidence with appropriate knowledge, capable of counteracting our natural inclinations, when these have a significant chance of leading us astray.** So it's a question of:

- learning to recognise situations in which, because of our natural cognitive functioning, we are more likely to make mistakes - in the sense of giving too much or too little weight to information in decision-making, of not judging information on the basis of its quality but of influences of various kinds other than the objective quality of the evidence and the source;
- learning to use in a fluid, advanced way, more powerful or more adapted strategies than the natural ones such as: strategies that use statistical tools, probabilities; the use of tools of social cooperation in order to put under control, in a voluntary way, biases such as the confirmation one.

3.5 Summary and general remarks

The mechanisms of epistemic vigilance, metacognitive sensitivity, awareness, and control over common biases are natural cognitive foundations of CT, as the **ability to assess the epistemic quality of available information and to calibrate one's confidence in that information for decision-making. These mechanisms can be defeated in complex, new (evolutionarily) situations, or simply by virtue of their intrinsic limitations.**

The CT naturalization effort should not be seen only in theory. **Building on a better understanding of cognitive processes is fundamental in the creation of evidence-based educational interventions, the next step being naturally that of validation by experimental testing first on a small and then on a large scale.**

The definition of CT and the investigation of its cognitive bases have thus allowed us to give practical indications for education, specific to each identified cognitive *building block*. We can now put forward some general considerations concerning the education of CT.

In terms of pedagogical engineering, the approach proposed here aims to **overcome the natural limitations of information evaluation capacities (first and second hand) and to extend their scope**. It is not a question of building an ideal critical thinker *from scratch*, but of understanding what - in the functioning of our epistemic vigilance mechanisms - makes it difficult for us to evaluate the information in circulation. We are therefore talking about a **tool-based CT, as opposed** to a **natural CT,** which would be part of our cognitive baggage. Equipping the natural CT allows access to increasingly **advanced** levels of CT. **Expert** CT includes tools that are specific to areas of expertise: for example, a person with expert CT in medicine is able to evaluate specialized information in that field, to know the most reliable sources in that field, and so on.

The educational question is then: how to equip natural CE?

Improving the evaluation of the epistemic quality of information requires the transmission of criteria and knowledge allowing an evaluation more adapted to complex situations. Tools for advanced CT therefore include:

- **advanced criteria for evaluating evidence and sources, knowledge to improve the assessment of the relevance and plausibility of information** ;
- **strategies and knowledge to better calibrate metacognitive sensitivity and reduce the impact of metacognitive bias**. For example, the improvement of domain knowledge, the use of external feedback, the explicitness of confidence assessment so that only indirect cues are used silently;
- **awareness of cognitive influences that may lead us to give less accurate assessments of available information, when these influences represent a risk**, how to counteract them using cultural (e.g. statistical) or artificial (e.g. the creation of cooperative situations) strategies and tools.

We therefore advocate, that CT education should aim :

- to build on existing capabilities by clearly identifying their limitations, including the conditions under which these capabilities become underperforming in relation to established objectives. This means, for example, adopting strategies to identify the origin of information, criteria for correctly gauging expertise while taking into account the fact that certain spontaneously used indices may take on a different value in new contexts, and criteria for understanding how knowledge is produced;
- to clearly identify the knowledge that enables natural capacities to be equipped in order to make them more efficient in contexts where they may become underperforming;

- to equip learners with this knowledge, helping them to appropriate it;
- to prepare learners to recognize contexts that may be challenging and to transfer their tools to these new contexts. Since it is difficult to self-assess, external assessments are aids to calibrating confidence. Upstream assessments (diagnostic assessments) can be useful to make the learner aware of his or her own limitations and thus motivate the search for better criteria.

Encouraging slower thinking and more controlled functioning may seem an obvious lever for CT education. However, taking the time or thinking more and looking for more arguments to justify a position does not always guarantee that the right arguments will be identified. Inhibiting one's initial intuitions in a general way is not a strategy to pursue, especially as these may be correct. Similarly, getting more information can become an endless exercise if we don't know when to stop.

The ultimate goal remains to enable the subject to better calibrate his confidence in the information at his disposal and in the decisions that flow from it. It is not a question of generally becoming more distrustful or wanting to check everything for oneself.

In contrast to these educational modalities based on "generalist" indications, we propose to equip CT with more appropriate criteria for evaluating the information and calibrating the confidence we have in this information. It is also a question of measuring the degree of success of educational actions on CT by measuring the adjustment of the calibration of confidence in relation to the available information.

It should be noted that having domain knowledge seems to be a key to better evaluate content and sources of information, as well as to better calibrate one's confidence in the domain in question. This consideration advocates a CT education that is not decoupled from the acquisition of rich and deep knowledge content. In practice, it would not be a matter of teaching to think well, but of teaching in such a way as to be able to think well, including through a knowledge base. Some knowledge acquired in the field in question can be immediately transferable to other fields, at the cost of adaptation and effective teaching strategies.

But how can transfer from one area of CT practice to another be encouraged? It is to this question that we will seek to answer in the section dedicated to the educability of CT.

4. Educability of CT

4.1 Educating general capacities

Is it possible to learn to think (more critically)? This question mobilizes a long debated idea about the very existence of general abilities such as thinking, problem solving, being creative... Can we talk about general abilities that, once acquired, will be applied to different and unrelated contexts and contents? Is it possible to learn to think? Don't we rather learn to think about given contents in a given context? And if we only learn in one context, do we know how to get rid of it in order to transfer the acquired capacity or knowledge to another context, more or less distant? These questions have sometimes been answered in opposite ways. As a result, the opportunity to teach general skills such as critical thinking has been alternately affirmed, denied, and then reaffirmed, with an alternating emphasis on local knowledge, to be preferred (or not) to general knowledge. Transferring knowledge and learned capacities is possible but has its requirements and constraints. Far from constituting a problem, this does not constitute a problem, but on the contrary provides practical indications on how to conceive an education for critical thinking that is transferable to new contexts, especially in everyday life.

Among the interventions aimed at developing critical thinking skills, those that prove effective - over time and independently of the initial learning context - meet these same criteria. This gives a first answer to the question whether we can learn to think more critically: yes, but specific conditions must be met for this learning to be usable later.

4.2. Is CT a general capacity?

Once upon a time, there was a Head of State from a faraway land. The chief felt that his country would soon be threatened by aggression from neighbouring nations that were far better endowed than his own militarily. It was necessary to be smarter and to seek to win the battle by intelligence rather than by force. Fortunately, the country could count on the world's greatest chess player! The Head of State was satisfied. Here was the solution: recruit this

Chess Master, teach him the rudiments of politics and military techniques and, thanks to his genius, defeat the enemy.

What do you think? Did the Head of State have a good idea, because the (general) abilities of the Chess Master will be automatically transposed to the battlefield and more globally to real life? Is this attempt doomed to disaster, because the capacities acquired by becoming a Chess Master are local and contextualized to moves of pieces on a chessboard? Is the idea promising, but not thought enough to lead to something?

This thought experiment is placed at the very beginning of an article in educational psychology that has become a classic on issues of transfer of learning, written by Harvard professors in the 1980s (Salomon & Perkins 1988). The issue of general abilities and their transfer had already had a long history when the article was written (in the late 1980s). This history crosses both the field of education and the cognitive sciences. At the beginning of the 20th century, the question was posed by Edward Thorndike, the father of behavioural psychology, who was very interested in education. It was in this field that he carried out his research on transfer, to understand whether we can consider learning Latin and mathematics as propaedeutic to other learning and abilities. In the 1950s, the question became pressing, particularly with the development of research programmes on artificial intelligence, the initial objective of which was to reproduce intelligent systems capable of solving problems or "reasoning" about any kind of content on the basis of a limited number of general heuristics that would apply independently of the occasional context chosen.

4.2.1 The problem of intelligence does not arise

The issue of general abilities also intersects with another area of research and long debate: that concerning the intelligence quotient and the g-factor or general intelligence introduced by Spearman in 1904. This idea responds to a fairly widespread intuition: "*The notion of general intelligence is based on a fairly shared intuition, according to which it is easy to distinguish between individuals who everyone qualifies as intelligent and those who are much less intelligent. However, a closer look reveals that abilities and talents can be multifaceted; one who excels in handling the subtleties of language may be less good at abstract reasoning,* *while another brilliant mathematician may be unable to manage his or her daily life. Nevertheless, observation of these individual differences does not detract from the intuition of a general form of intelligence that would apply to many areas of life. In fact, data collected over centuries of intelligence testing - a comprehensive test usually includes several individual subtests - support this intuition.*" (Ramus, 2012). The general public, and teachers in particular, are familiar with a critique by Howard Gardner of Harvard University of the idea of general intelligence and a score, the IQ, that quantifies it. According to Gardner, IQ tests do not measure abilities such as social skills that play an important role in everyday life. While this objection is correct, it does not detract from the fact that IQ measurement is reliable (performance on all tests and subtests is correlated) and reproducible (retaking the test does not change the result), and that it is correlated with effects such as academic success, which it can predict. But is there a particular cognitive function, a specific mechanism at work in our brain, that would explain g and the correlations between IQ test scores? The answer is more complicated and allows us to eliminate a first suspect in the question of general abilities. In fact, no study has been able to find a specific link between a particular faculty (among the possibilities: attention, speed of information processing, working memory) and g or IQ. On the other hand, it is obvious that each test simultaneously mobilizes a multitude of cognitive functions.

This is also true outside of testing: cognitive functions are highly specialized, but they work in an intertwined way to contribute to different facets of our behaviours. The answer to the apparent existence of a single basis for intelligence is therefore this intertwining and interdependence of different cognitive functions. *"What we call general intelligence is simply an emergent property, resulting from the cascade of environmental, genetic, brain and cognitive factors that influence performance on different tests*." (Ramus, 2012). Thus, although different talents exist, the statistical correlation between our different abilities remains real, because they often operate and develop together by influencing each other during our development, even though they may draw on different aspects of brain maturation that are influenced in common by genes and environment. However, the problem of intelligence does not really arise when it comes to general *versus* local ability to think. Rather, the issue is about the heuristics we mobilize to solve different types of problems, to think about different types of content. Of course, g could (eventually) influence the quality of these heuristics, but has nothing to say about their more or less broad applicability. So let's focus on heuristics.

4.2.2 General heiristics: yes or no?

A heuristic is a rule, a strategy, simple, to tackle a problem. The 1950s-1970s were characterized by a certain confidence in the existence of general heuristics (effectively underpinned by factor g) to address broad classes of problems and other intellectual challenges. An example of a problem-solving heuristic is to divide the problem into sub-problems and then solve one problem at a time. Another example is to represent a problem in different aspects. And so on and so forth. During the golden age of heuristics, this type of approach gave rise to artificial intelligence programs with the illustrative name "*General Problem Solver*" (GPS), created by two of the fathers of cognitive science, Allen Newell and Herbert Simon, along with J.P. Shaw. The heuristic at the heart of GPS (not to be confused with the *Global Positioning System,* which displays a completely different type of intelligence) is that of *means-end analysis*. A problem is defined by its initial and final state; each step in solving the problem puts in place a means to make the initial state more like the final one. After this operation, the two are compared and a new procedure starts again to reduce the difference later. And so on and so forth. General principles are therefore systematically applied, without questioning the specific nature of the problem to be addressed. Beyond AI, there is the vision that "good thinking" means having a set of good general heuristics: for problem solving, for memorizing, for inventing, for making decisions.

The corollary of this vision is that each local knowledge and knowledge database plays a secondary role in thinking. Of course, nobody denies that to become a chess master it is necessary to master the rules of the game or that a good doctor needs a knowledge base of physiology and pathology. But possessing and even mastering these local rules and databases is insufficient to explain the superior performance of Chess Masters and good doctors, which are rather to be reduced to their mastery of general heuristics. Within the framework of this vision, the Head of State has indeed only to provide his chess master with the appropriate knowledge base so that the latter, strong of his general heuristics, is transformed into a war machine. The problem of this vision is that it is not supported by the facts. Four fields of empirical research undermine more particularly its bases.

4.2.3 Strenghts of the experts

What if we analyze what chess masters really know how to do, beyond their ability to win games? During the same years that the idea of the "general thinker" was asserting itself, researchers studied specific aspects of the performance of experts, including chess masters. Several questions were then raised: do chess masters show a breathtaking memory? Are they able to keep the memory of entire matches far away in time? Does this attitude translate into a more global capacity to remember? Is it possible to train our memory in one area (chess or other) and see the benefits elsewhere? The results of the studies conducted at the time were rather disappointing.

For example, Chess Masters are able to learn by heart the staggering figure of 50,000 different chess configurations. However, their ability to learn by rote and to reproduce the position of chess pieces outside the rules of the chessboard is not superior to anyone else. In fact, the Chess Master's ability is superior only if the configurations to be learned make sense and relate to chess knowledge and rules (de Groot 1965; Chase & Simon 1973); if he has more time, he succeeds slightly better than novices (Gobet & Simon 1996). The vision, which emerges then from expertise goes in an opposite direction to that of general heuristics. *"People who have developed expertise in particular areas are, by definition, able to think effectively about problems in those areas. Understanding expertise is important because it provides insights into the nature of thinking and problem solving. Research shows that it is not simply general abilities, such as memory or intelligence, nor the use of general strategies that differentiate experts from novices. Instead, experts have acquired extensive knowledge that affects what they notice and how they organize, represent, and interpret information in the environment. This, in turn, affects their ability to remember, reason, and solve problems*." (Bransford et al. 2000).

The experts use a broad knowledge base of relatively specific knowledge and know-how. For example, a scientist by profession, a physicist, has a knowledge base of the laws and principles of physics (Chi, Feltovich, Glaser 1981). This knowledge is not just arranged somewhere or independent of each other: it is easily mobilized and is strongly intertwined with each other, forming a network of usable knowledge (Chi, et al. 1982).

Experts are able to quickly recognize the situations in which this knowledge applies, they perceive the "deep" structure of the knowledge. For example, faced with a problem concerning an inclined plane, a physicist will immediately see the laws of motion at work. The chess master, on his side, perceives the configurations on the chessboard with their meaning, not simply as disjointed pieces: he does not have to try to link them by an abstract process, but sees them directly as linked by a pattern (Bransford et al 2000[195]). Experts think forward: they use their knowledge and principles to integrate findings and new problems (Larkin et al. 1980). They seem to use metacognition: they assess their level of understanding and ask themselves whether the knowledge at their disposal is sufficient to solve the current problem. They may question what counts as expertise in their field (Bransford et al. 2000). Bransford et al. (2000) summarize the specificities of experts in six points[196]. Furthermore, they consider the notion of expert to be important for learning, as it helps to show what a learner can achieve in terms of thinking skills in a field.

The novice, who behaves in the opposite way to the expert, lacks a broad and deep knowledge base, lacks the habits that allow him to quickly recognize a deep structure, remains "stuck" to the surface of the problem and tends to reason first about the new problem and its unknowns, only to go afterwards to look for in his knowledge those that could help him to solve it. These two figures emerge from a wide variety of disciplinary fields: mathematics, programming, medicine, science, and of course chess. Thus a theory of expertise is born: experts "see" things differently; faced with the same problem, they have rapid access to interwoven knowledge in which a new piece of information finds its meaning or can be questioned in the light of general principles (Chi, Glaser, Farr 1988).

The notion of expertise also receives a new dimension in the field of artificial intelligence where, from the 1970s onwards, "expert systems" begin to supplant "general problem solvers". In medicine, for example, researchers are noticing the powerful effects produced by broad knowledge bases. They began to regard general heuristics as 'weak methods', which cost little in terms of the database but also have weak results. The new expert systems aim to simulate the real capabilities of experts, based on the characteristics described above and thus on massive databases on, for example, different medical diagnoses.

4.2.4 The brain is not a muscle

Other research is helping to dispel the myth of all-encompassing general ability training. In particular, they are debunking the myth that the brain is some kind of muscle that we can train to improve our overall abilities, regardless of context and content. This is the case with short-term memory.

Researchers have asked a subject to memorize as many numbers as possible and then repeat them. This task is reminiscent of memorizing the digits of a telephone number before writing it down somewhere. On a daily basis, we all realize that this capacity is limited (around seven items). This feeling is confirmed by research conducted at the dawn of cognitive science (Miller 1956). It is also shared by the subject of the experience we are going to relate. Except that, after two years (at the rate of two to five days of training per week, two hours per day), the subject in question - known by his initials SF - is able to recite 81 numbers in a row (Chase & Ericsson 1982, Ericsson & Chase, 1982). To make a long story short, his lengthy and expensive training did not actually develop his short-term memory as an overall cognitive faculty, but only his ability to memorize numbers. For the same memorization task but with letters, his performance drops to seven memorized items! When asked about his training, SF revealed that he used strategies to compact several numbers into a single mental "object" (a "*chunk"s*, is the technical term used): fascinated by the Olympic Games, he associated sets of numbers with an event, a race for example, and thus reduced the space in his short-term memory.

This study is often used to support the idea that the ability to integrate new information into our mental space depends on the knowledge we have and our ability to use it wisely in thinking. In reality, to think, we use our short-term working memory - the space in which we mentally manipulate information and representations - but also our long-term memory.

4.2.5 Knowing and know how to apply

It is trivial to point out that we are always thinking about something, about a certain content, and that this content belongs to a specific area. Thus, does knowledge in a domain influence our performance, understanding for example (Recht & Leslie 1988; Willingham 2007, 2019)? Numerous studies show that prior knowledge (the level of expertise) is a predictor of performance differences between subjects, regardless of the developmental factor (age of the subjects[197]). However, it is less trivial to show that different reasoning processes are involved depending on the domain. As an example of the influence of content on thinking skills, Wason's selection task: "*Four cards with a number on one side and a letter on the other are laid out flat on a table. Only one side of each card is visible. The visible sides are as follows: D, 7, 5, K. Which card(s) must you turn over to determine the truthfulness of the following rule: If a card has a D on one side, then it has a 5 on the other side? Do not flip a card over unnecessarily, or forget to flip a card.* "(Wikipedia: Wason's selection task). Fewer than 25% of the people submitted to this test give the correct answer (D & 7). Furthermore, formal training in logical abilities does not appear to have an impact on test performance. However, performance tends to improve significantly if the same task is presented under a "moral aspect", so the content changes but not the applied logic. *"Four people are drinking in a bar and you have the following information: the first person is drinking an alcoholic drink, the second is under 18, the third is over 18 and the last is drinking a non-alcoholic drink. Which person(s) should you ask about their age or the contents of their drink to make sure that everyone follows the rule: if a person drinks alcohol, they must be over 18.* "When identification of the logical error reveals a cheat, performance improves and success in resolution rises to 85% (Cosmides & Tooby 1992)!

Let us now take the case of a person who has learned that it is important, when faced with a problem, a choice, to operate, to use particular metacognitive strategies: this person would have taken a course in critical thinking. In the presence of a new problem or circumstance, he or she will know that it is important not to give in to confirmation bias, or to the argument of authority, that it is important to suspend judgment in order to fairly evaluate different options, to doubt, to question whether the argument is valid and whether it is also supported by facts (thus possibly true). We can also know that, when faced with a complex problem, an effective strategy is to divide it into several smaller sub-problems. But which ones? How do we operate the division? Up to what level of division should we go? In order to answer these practical or theoretical questions, it is useful to have some experience with similar problems and knowledge about the constituents of the problem. The importance of content for the acquisition of higher thinking skills is therefore not limited to the fact that thinking needs knowledge content to be exercised: this capacity improves because of the amount of knowledge possessed and the more we know about a certain area, the better we are able to think about it effectively and critically.

The problem of general heuristics, disconnected from a domain of application and domain knowledge, is therefore that of their application. Even if we follow the right rules to solve a problem, if we don't know when and how specifically to apply them, it doesn't work. The studies on expertise and those on short-term memory revalue the importance, in reasoning, of knowledge (the database) and, in the ability to think, of memory, as a function that allows this knowledge to be stored. Thinking effectively, understanding in depth, finding an expert solution are all abilities that are highly dependent on domain-specific abilities and knowledge content.

4.2.6 Strategies (learning to learn)

Another part of the literature in cognitive and educational psychology useful for our discussion refers to the concepts of strategy and learning to learn (Fayol & Monteil 1994). The notion of strategy is particularly developed in the literature on memory, and memorization strategies: rereading, rehearsing, preparing summaries are examples (Bjorklund & Harnishfeger 1990). There is now a vast literature of scientific studies concerning the most effective techniques for memorizing knowledge - the practice of testing, spaced learning, seeking *feedback* and, more generally, active and strenuous engagement in learning (Brown, Roediger & McDaniel 2014). A strategy has specific characteristics (which are similar to heuristics as described above, at least the general strategies[198]). Strategies differ from basic procedures and algorithms in that the subject

- has multiple procedures to choose from;
- it selects procedures on the basis of criteria such as purpose, circumstances, knowledge of its possibilities.

Thus, a strategy is not an "automatic" response to a given situation, imposed by the fact that it is the only one available or the only one evoked by the situation. Moreover, it is not put into practice automatically, according to a pre-constituted scheme of steps. The subject makes a choice and then guides his or her action and evaluates it in a way that requires attention. Implementing a strategy is therefore cognitively costly, but the gain is that of adaptability and flexibility in relation to circumstances and subjects. The ability to implement a strategy is therefore part of high-level thinking skills, such as critical thinking (Resnick 1987). A fairly general observation is that strategies, once acquired, are not systematically reused. For example, summarizing to learn better is a strategy that can be taught; students use it during the lesson but abandon it in the next session (Fayol & Monteil 1994).

A series of difficulties are interposed between the fact of having learned a general strategy and that of using it in life - even at school: the difficulty of perceiving the added value of the strategy, in the absence of appropriate feedback and information on the relevance of its use (already that we have to conceive that what we do can have an impact on our success, that it is not a kind of destiny but depends on effort); the cognitive cost of effort; finally, the role of prior knowledge[199]. So much so that the learning of procedures would be facilitated when their introduction and the training to use them would be done on contents mastered, at the level of knowledge, by the subject. Prior knowledge can thus constitute an obstacle but also an asset for the acquisition of strategies because it avoids cognitive overload.

4.3. Learnng to think (more) critically: a false hope?

There is no consensus that CT interventions are effective. Why? There are two main reasons for this absence:

- different authors use different criteria to define what constitutes a positive result;
- the authors do not use the same CT definition.

These two difficulties are compounded by the fact that the interventions evaluated are very different from one another: in terms of duration, typology of activities, but also in terms of objective and theoretical vision of what CT is. All of these difficulties make quantitative meta-analyses and qualitative systematic reviews of the literature difficult to conduct and reduce their relevance.

4.3.1 Daniel Willingham's pessimism

Cognitive psychologist Daniel Willingham has conducted a brief analysis of the effectiveness of the best-known programs for teaching critical thinking *per se* (Willingham 2007). He synthesizes the common characteristics of the more widespread methods in three points:

- they presume the existence of skills that can be practised independently of context and content, so that teaching takes place outside a disciplinary framework;

- some are of long duration (three years, several hours of instruction per week);
- they all use examples of critical thinking and then ask to apply; some use abstract problems (*Ruven Feuerstein Instrumental Enrichment*), others use mystery stories (*Martin Covington Productive Thinking*), group discussion of everyday problems (*Edward De Bono Cognitive Research Trust*: CORT).

Studies that measure the effects of these interventions have several limitations:

- students are evaluated only once, so we can't know if the effects last;
- there is no control group, or the control group does not carry out an alternative activity (passive group) ;
- there is no transfer measure to real situations or situations different from those used in the investigation ;
- only a small proportion of these studies have gone through the scientific process of publication and peer review;
- when we talk about the positive effects of these methods, we are not necessarily talking about the same thing or about something specifically related to critical thinking[200].

Despite all these difficulties and limitations, reviews of the literature or texts on critical thinking often indicate that educational interventions for critical thinking have positive effects, a conclusion that Willingham rejects[201]. Drawing on the literature cited above, Willingham (2007) also rejects the idea that thinking can be considered a *skill* similar to cycling. Thinking is highly dependent on the content of knowledge. Reminding someone to use his or her critical mind, or to think about looking for alternative hypotheses, etc., is similar to the situation where we tell someone: "Be careful"[202]. This type of exhortation can serve as a "reminder" but actuually that someone doesn't necessarily know what to pay attention to and how to avoid the undesirable behavior. Moreover, when we seek to understand a new problem, we use the knowledge stored in our memory as well as the context. This makes understanding faster, but anchors it to the contextual content. These considerations lead Daniel Willingham to affirm that teaching - generally speaking - to think or think critically is impossible[203].

Is there no hope of transferring gains from one context to another - including "general" gains such as the ability to solve a certain type of problem? Willingham points to two: familiarity with deep content and knowing that one has to look at that deep content[204]. Familiarity depends on repeated, automated practice[205]. The second strategy is metacognitive: thinking about looking in memory to see if there is something similar[206]. When this does not come automatically through familiarity, one must at least have in mind the principle of doing so. To really implement it, knowledge and practice are necessary. Without that, we know what we should do but we don't know how to do it. We can conclude that the ability to think more critically depends strongly on content.

This is also true for more specialized forms of "expert" thinking, such as scientific thinking. Even in the case of scientific thinking, success depends not only on knowing the procedures and strategies, but also on knowing when and how to apply them[207]. The final conclusion is that although we have natural abilities, such as the ability to reason about causes and conditional probabilities (an ability that is intuitively limited from early childhood), we can continue to make mistakes in the use of conditional probabilities and causal reasoning even in the presence of more sophisticated abilities and knowledge[208]. Critical thinking instruction thus relies, in part, on teaching students strategies for "better thinking", and in large part on how and when to deploy these strategies[209].

4.3.2 Diane Halpern's optimism

Diane Halpern, also a cognitive psychologist, is much more optimistic about the possibility of educating critical thinking and the success of existing methods (Halpern 2013[210]). She cites studies conducted by other researchers or groups, such as a meta-analysis by the *Thinking Skills review Group* (Higgins et al. 2005). It includes interventions that involve metacognition on learning and thinking processes, activities related to creativity but also reasoning skills (verbal, spatial, logical).

Halpern also cites studies that have evaluated the impact of specific interventions on thinking skills, both large and small scale. A national evaluation in Venezuela, conducted by the government with the assistance of Harvard on students averaging 13 years of age, measured the effects of a large-scale, randomized, controlled intervention (Herrnstein et al. 1986). Van Gelder (2001) has evaluated his own program focusing on learning argumentation strategies, a program dedicated to university students. Marin and Halpern (2010) analyze the incremental impact of explicit and implicit CT skill interventions that can be measured by the Halpern

Critical Thinking Assessment (HCTA): the target of the intervention and assessment are adolescent students in upper secondary school. The results are comparative: explicit instruction works better than implicit instruction (Heldsdingen et al. 2010). There is also a reality test in the Netherlands of one intervention compared with no intervention. Also cited are: Facione (1991); Facione (2000); Lehman, Lempert, Nisbett (1998) who test their interventions *via* multiple choice tests and mini-scenarios. For Halpern, the most significant group of studies, particularly for transfer issues, are those conducted by Nisbett and reported in Nisbett, Jepson, Krantz (1993), Nisbett et al. (1993), Nisbett et al. (1987) and Fong, Krantz, Nisbett (1986). See also: Nisbett 2013).

Nisbett's interventions focus in particular on aspects such as logic, causal and probabilistic reasoning, and cost-benefit analysis. Nisbett uses realistic situations both in his critical thinking pedagogy and in his testing. The most "extreme" view of the ability to measure transfer to real-life situations is that conducted by telephone without the subjects of the experiment knowing that they are taking a critical thinking test. Based on his results, Nisbett considers that critical thinking is a skill that can be learned and transferred to new contexts and content. Kosonen and Winne (1995) show positive results for learning rules of reasoning applied in different contexts. Halpern's conclusion is that transfer is possible modulo this type of strategy: repetition in several contexts and teaching the rules. She considers that successful teaching of critical thinking should point the way to more general teaching when we want to achieve transfer. On the contrary, instruction not specifically dedicated to improving critical thinking does not seem to have any effect on it[211]. Henceforth, Halpern's process for improving critical thinking skills is based on four educational strategies:

- teach critical thinking in an assertive way, without expecting it to develop on its own as part of "normal" instruction;
- explicitly teach strategies ;
- practice the same strategies in a variety of contexts and on a variety of contents;
- use pedagogical tools that have proven useful for other learning, notably the "test" strategy: testing oneself to better memorize, at the end as well as in the middle of learning[212].

This strategy can be used to multiply practice in various contexts and with different contents. Despite the difference between the pessimism of one and the optimism of the other, these four indications are common to Willingham and Halpern, who in any case share the same field of knowledge about learning and transfer.

However, Halpern may be overly optimistic. Her vision implies not only the improvement of some of our capacities (language analysis, use of probabilities, plausibility, understanding of what counts as evidence...) but also the establishment of capacities of "resistance" to our natural biases and tendencies, in "natural" conditions and contexts that involve limited time, stress, emotions, social motivations[213]. Moreove, the studies cited by Halpern (2013) actually provide weak evidence of the educability of CT using existing methods.

Let us look at some of them in more detail. We will focus on the meta-analysis by Higgins et al (2005), the largest experimental study conducted (Herrnstein et al. 1986), the study conducted by Halpern and herself (Marin & Halpern 2010) and the studies by Richard Nisbett, which seem to be the most promising in terms of transfer.

4.3.2.1 A meta-analysis

The meta-analysis, quantitative, analysis published in 2005 by Higgins et al. is one of the systematic impact studies of the EPPI Centre at UCL. The EPPI Centre plays a role in promoting *evidence-based* education and, more generally, in the field of social policy. It also serves as a study production and meta-analysis body for the European EIPPEE. We take the opportunity of this study to emphasise that meta-analyses are always to be favoured in impact assessment and that the motivation for the analysis is that CT is a pedagogical objective in curricula in England and Wales. Higgins et al (2005) consider not only CT, but also thinking, planning skills, and metacognition in a general sense[214]. Thus, their review of the literature includes programmes such as *instrumental enrichment* (Feuerstein et al., 1980), *philosophy for children* (Lipman et al., 1980), *cognitive acceleration through science education (*CASE) (Adey et al., 1989), *Somerset thinking skills* (Blagg 1988) and programmes using methods centred on dialogue, argumentation, etc. (Blagg 1988). The meta-analysis included 29 studies, selected following the criteria of the EPPI Centre. However, these criteria are quite broad (in fact, only 13 were randomized[215]). The results of the meta-analysis, considered positive by the authors, refer to measures of impact on various cognitive abilities, some of which do not belong to the domain of critical thinking (e.g. Raven's matrices: effect size 0.62, typically included in IQ tests but with no specific relationship to critical thinking or other forms of information analysis). The authors also indicate an improvement in academic performance, which can be related to programmes specifically aimed at science skills, such as CASE, but difficult to relate to interventions to improve thinking (academic performance is highly dependent on knowledge). There is great heterogeneity between interventions (some relate to how to foster optimistic ways of thinking). Basically, these findings are more about teaching methods than about teaching critical thinking - and their results are indeed consistent with those obtained by Hattie and Marzano on methods involving metacognition (Hattie, Biggs & Purdie 1996; Hattie & Donoghue 2016; Marzano 1998). The hypothesis that they might ultimately support is therefore that positive results can be achieved when we address the education of thinking skills and not disciplinary content, but not in the direction of CT. We also find a great deal of heterogeneity in the age of the subjects in the different studies and in the types of measures (standardized or non-standardized tests). It is therefore preferable to be very careful with the interpretation of these results. The meta-analysis also highlights the existence of publication bias, which suggests that studies published with negative results are biased.

4.3.2.2 A large intervention

Herrnstein et al (1986) describe the results of an intervention on about 400 college students (13 years old) in Venezuela. This intervention is unique in that it did not focus on curricular content but on skills such as 'scientific' reasoning (observation and classification), critical use of language, argumentation, creativity, problem solving and decision-making. The intervention consists of a total of 100 lessons of which about 60 are used in the classroom. The students are all of LSES origin. The intervention lasts one year, and is systematic: three lessons take place over the four school days, the fourth being used for revision. The lessons are given by external teachers selected and prepared for the task. A battery of tests is administered to students in test and control classes (passive control) in pre-test and post-test. The battery consists of standardized tests and not for a total of about four and a half hours of taking. Non-standardized tests measure the abilities worked on by changing content. Students in the experimental group are better in all pre-tests than students in the control group. The overall effect is considered positive because students in the experimental group make more progress, especially in the non-standardized tests, than students in the control group. Herrnstein et al. (1986) correctly pose the question of what can be considered a success for this type of initiative and how to distinguish between effects produced by the contents of the intervention and effects due to the "intervention" or Hawthorne effect. Indeed, they notice an institutional and teaching effect.

4.3.2.3 A measure of impact for a specific intervention

Marin and Halpern (2010) evaluated the impact on graduate students of an explicit and practical type of intervention in line with Halpern's four pillars of critical thinking instruction: working on attitudes and encouraging students to reflexivity; teaching and practicing critical thinking skills; training for transfer, identifying concrete situations in which to adopt the strategies learned; metacognition as reflection on one's thought processes. The intervention is compared to an implicit intervention (teaching a cognitive psychology course with common but implicit contents) and with a passive group "waiting" for its course. Respectively, the study involved 28 + 18 + 24 students. The intervention consisted of a pre-test, four lessons, and then a post-test. It took place over three weeks, at the rate of two lessons of two and a half hours per week. Participation was voluntary and took place after class or on Saturdays. Each lesson was attended by an instructor, two research assistants, and a technician. The intervention was based on computer support that the student used alone and equipment for use in class with the instructor. Both groups, explicit and implicit, progressed between pre-test and post-test, but the experimental group made the most progress.
To overcome the limitations of the study (small numbers, self-selection bias, etc.), the study was redone with six lessons given directly during school hours, twice a week for six weeks (40 + 38 + 30 students for the experimental group, active control, passive control). The passive group did not conduct a post-test after some time. The results show performance gains for students in the explicit group but not for those in the implicit group.

4.3.2.4 Measuring distant transfer

The psychologist Richard Nisbett is in a diametrically opposite position to Willingham regarding the ability to learn the general abstract capabilities of a particular content. He argues that his view has been modified after undertaking studies with Fong and Kranz to demonstrate precisely this limitation. In studying the impact of instruction in statistics, he would have realized that, in reality, this type of instruction has general effects that go beyond the content taught. Especially since the effects would already be visible with relatively short training sessions (Nisbett 1993). Nisbett tested its effects in ecological situations: by calling the subjects at home and pretending to conduct a survey, he asked questions about statistical skills and their application in cases different from those instructed. However, the questions asked could be reminiscent of the concepts of statistics being worked on, as they involved the language of mathematics. In addition, participants in the statistics course answered the survey questions shortly after completing the course. Nisbett has replicated its studies on other types of content, such as rules for assessing causality, cost-benefit analysis, etc., in the course. He believes that in areas where we have intuitive (and imperfect) rules of resolution, it is possible to teach more sophisticated systems of rules (the rules of statistics, economics) and the learner can then apply them generally. However, Nisbett acknowledges that differences in performance exist in the transfer of a general rule to concrete situations. These differences can be attributed to the transparency of the problem. Indeed, some problems allow the system of rules needed to solve them to be seen through their superficial content; others do not and therefore applying the rules becomes difficult in these cases. Nisbett also acknowledges that there are relatively few systems of rules that allow for this kind of general treatment: rules that identify the good cause, contractual patterns in the social sphere, statistical rules such as the law of large numbers[216].

4.3.3 A new synthesis: Abrami et al. (2015)

Recently, Abrami et al (2015) undertook to synthesize the various studies on the impact of methods for teaching critical thinking. Their research was carried out in 2003 and updated several times (most recently in 2009). A total of 684 published studies were selected for the final analysis because they met the relevance criteria established for the meta-analysis.
Such an undertaking naturally comes up against several difficulties:

- unclear and non-consensual definitions of CT;

- a wide variety of educational methods, in particular because they follow different points of view on the general or rather specific (content) nature of thinking skills, and because they are of varying duration and intensity;
- the multiplication of methods for the assessments of success, conducted using a variety of instruments, some standardized - such as the Watson *Glaser Critical Thinking Assessment*, the *Cornell Critical Thinking Test*, the *California Critical Thinking test*, and the *Critical Thinking Disposition Inventory* -, others not. Among the latter, some forms of assessment are created by teachers, by the researchers who designed the intervention, or by other researchers, or they are unthought-out scales and tests to assess critical thinking (including academic performance);
- the often quasi-experimental or pre-experimental nature of the available studies: when edicational methods are tested in "ecological" (classroom) contexts, they often lack control groups or do not perform random distribution of subjects. In practice, many published studies are limited to comparing pre-test (pre-intervention) and post-test (post-intervention) results for the same group of subjects, making it impossible to control for test effects, master effects and the effects of other unidentified variables on the final outcome.

However, the authors of the meta-analysis decided to also include these quasi-experimental studies in their selection by considering the pre-test result as equivalent to that of a control group. It must therefore be recognized that the current results, taken in their complex, are not very meaningful and can hardly enlighten us as to the effectiveness of teaching critical thinking as a general or specific capacity.

With these considerations in mind, do the results of research on critical thinking education show, in their complexity, positive effects? The answer is positive, with several caveats that stem directly from the difficulties discussed above. It appears that the different methods examined make it possible to develop one or more abilities identified as part of critical thinking, either general or content-specific, depending on the objectives of the study and the method adopted. This is true at all educational levels, across all disciplines and for all the teaching strategies adopted.

Teaching strategies are classified into four main groups, according to a classification proposed by Ennis (1989): general education (principles, skills, dispositions are taught in a general way and not in relation to any particular content); immersive teaching; infusion teaching; mixed teaching. (In both immersive and infusion teaching, content is important, but in the former, the general principles of critical thinking are still taught explicitly as course objectives). Other pedagogical modalities are considered separately, including: individual study or teaching; teaching through dialogue, group discussion, formal debate, and other collaborative methods; presentation of authentic problems to be solved, grounded in rich content; coaching by a tutor. Also compared are short, medium, and long term interventions. All strategies, methods and durations give positive results compared to control groups without critical thinking instruction (when present) or pre-tests before critical thinking instruction. However, the impacts (effect size) are rather small. Some modalities of instruction are more effective than others for the development of 'general' skills (not linked to specific subject content): on the one hand, teacher-guided discussions and, on the other, exposure to authentic situations and rich, situated contexts with examples, especially when two methods are used: problem solving and role playing. These two approaches seem particularly effective when combined and when some form of tutoring is added. Studies in which the three modalities are combined show significantly larger effect sizes than the three individually (tutoring alone, however, does not show a larger positive effect than the other strategies and methods mentioned above).

Naturally, several questions remain open in the light of these results. The fact that any intervention has an effect only confirms that "something is better than nothing". But we cannot know whether the effect is sustained or whether it is transferable beyond the tests carried out, and therefore whether it has a real impact on the lives and attitudes of those who have received some form of instruction in critical thinking. In practice, the studies reviewed make it possible to establish that the various forms of instruction that have been tested and whose results have been published make it possible to produce learning that is at best a precursor to generative and transferable critical thinking. This is encouraging, but invites even more to ask the question of transferability in any form of critical thinking instruction. It should not be forgotten that some methods seem - at this level - more effective than others, and that they therefore provide indications on the methods that should be developed and tested as a matter of priority: presence of authentic arguments, contents and examples, accompaniment on the part of the teacher in the form of tutoring.

4.4. Difficult does not mean impossible: reality and difficulties of transfer

If skills such as thinking or problem solving are general, then it should be easy to transfer them from one learning context and content to another - or even happen automatically. We have quoted Edward Thorndike who, in the early nineteenth century, set out to test in schools whether learning Latin and mathematics had a measurable impact on other cognitive functions. His answer was negative. Further proof that the brain is not a muscle, and that mental faculties are not automatically developed by specific learning. This impression is confirmed by many studies on learning transfer, which can be summarized as follows: transfer is the beast of learning (Salomon & Perkins 1989). However, intuitions and everyday experiences seem to suggest that the image of learning as completely dependent on context and content (not only in the acquisition phase, but also in the reuse phase) is excessively negative. We are learning to ride a bike and we are not dependent on our first bike for life. We learn to cook vegetables and less effort is required to learn how to cook fish. We learn to read and even we don't have the ability to understand everything, because of our knowledge, reading as a tool is generalized from one context to another. We learn a way of conducting experiments or of recognising a certain typology of fallacious reasoning, and we have the impression that we are capable of conducting experiments in other more or less distant fields and of reusing these typologies from contents that are quite distant from each other.

What does an expert do when faced with an atypical problem? He can draw on his database of knowledge and know-how. He can recognize the deep structure of the problem. But if the situation is really new, this will not be enough. Several studies show that when faced with a new problem, scientists use analogical reasoning. In this way, they seek to bring novelty back into familiar territory by trying to identify elements in the structure of the former that bring it closer to problems they can deal with. Some studies therefore suggest that, in this type of challenge, the expert is forced to fall back on general heuristics - such as the use of reasoning by analogy - but without forgetting his specific knowledge. And that he does so in order to transfer his skills to a distant field.

An even more general consideration leads us to doubt the impossibility of transfer. Basically, we find ourselves almost constantly confronted with new situations - or at least with situations that present both many similarities with those encountered in the past, but also inevitably some differences. If we could not transfer our learning and know-how at all, it would be impossible to cope with the new. Not only do we respond to novelty, but we often do so, wrongly, by ignoring its novelty and transferring inappropriate solutions as if we were facing a known situation. In other words, we often apply strategies available to us, blindly and sometimes unconsciously, to a new situation. So why does the transfer sometimes not take place? The question of transfer is probably a matter of degree, quality, and method.

4.4.1 The concept of transfer is fundamental for learning

The distinction between transfer and learning is blurred, of course, because we cannot consider learning as such if it does not manifest itself in a minimal level of transfer from one initial situation to a new one. The notion of transfer is therefore an essential part of the notion of learning, and we cannot set ourselves the goal of the latter without aiming at least partly at the former. The consideration of the centrality of transfer is all the more important when we ask ourselves the question of the objectives of education[217]. Education makes it possible to pass on to the younger generations the knowledge and skills developed by their predecessors, but one of the objectives is also to enable them to make use of this knowledge and skills, and possibly to improve them. For this, learning to transfer is fundamental. Learning to transfer is also necessary to fulfil another objective of education, that of developing thinking and reflection skills, to improve the ability of future citizens to decide, make rational choices, interact with others and with their environment in a flexible way. The question is therefore threefold: can we transfer skills, knowledge and attitudes relating to critical thinking, which would have been learned in a certain field, in a certain context, to another? To everyday life? In a flexible way?

When Thorndike sought to test the validity of the hypothesis of formative disciplines (a thesis dear to the psychologist Alfred Binet, inventor of the IQ test), he was actually thinking of another theory, his own (Thorndike & Woodworth 1901, a, b). According to him, the transfer between an initial, learning situation and a new application situation is possible, but only if there are common elements common to both situations (*common element theory*). In this general formulation, the theory is widely accepted. The classical and widespread distinction between near and far transfer goes back to this idea: transfer between situations which share identical elements, or which present many similarities, including surface similarities, is said to be near. For example, we can learn to draw in pencil and then transfer this skill to progressively more distant areas: chalk, watercolour, oil... Some skills can even be transferred to the photo. This type of transfer is more likely to succeed than the distant transfer - which is not guaranteed (Sala & Gobet 2017).

The problem with this theory is represented - as we shall see later - by the ambiguity of the term "element" and by the difficulty of defining in a single way the distance between elements, i.e. the type of similarity between situations allowing the transfer. In particular, attention has focused on the "internal" elements of the learning situations and target situations to be transferred to: objects, actions, events present in both, and the logical relationships between these internal elements (Gyck & Holyoak 1980, 1983; Gentner et al. 2003). In the 1980s, for example, Anderson took up the theory of element identity and argued that transfer depends on the existence of similarities between the knowledge and cognitive operations required by the learning task and the target task (Singley & Anderson 1989; Anderson, Reder, Simon 1996).

4.4.1.1 Transfer and analogy

A doctor is faced with a difficult case: he has to treat a patient who suffers from stomach cancer; the patient will die if the tumour is not destroyed, but the operation is impossible. He can use radiation powerful enough to get rid of the tumour, but its passage through healthy tissue will also destroy it. If maximum power is not used, the therapy is ineffective. How can you destroy the tumour with radiation without destroying the patient's healthy tissue at the same time?

This problem is known as the "*Dunker's radiation problem*", named after the psychologist who introduced it in 1945 as part of a problem-solving exercise (Dunker 1945). Presented in this way, the chances of solving the problem are rather low. Other psychologists, having taken it up again later, have estimated them in their studies at about 10% (Gick & Holyoak 1980, 1983). This figure may seem too low to us, because we have the impression that the problem is not that difficult. Our intuition is that we have become familiar with cancer treatments and that certain solutions come to mind spontaneously... because we know them. Knowledge of how radiation works is the key to solving this problem effectively (Helfenstein & Saariluoma

2006). Another factor that significantly increases the chances of success is the availability of an analogous solved case, as the following example shows.

A small nation is dominated by a dictator in a well-protected fortress. The fortress is located in the centre of the nation, surrounded by villages and farms, and several roads lead to it. A general decides to defy the dictator and attack the fortress. He knows that by mobilizing his entire army at once, he has a good chance of winning. So he gathers all his troops at the mouth of one of the roads leading to the fortress, but there he learns that the dictator has scattered mines along the roads so that only a few men can pass through at the same time without setting off an explosion. A large contingent of men passing along a road at the same time could only detonate the mines. Both the contingent and the villages surrounding the roads would be destroyed. The general came up with a plan: he divided his troops into small groups and launched them each on a different road. The men will all find themselves in front of the fortress, but without going through a single road. He launches the attack and emerges victorious.

We've got that story. We also know that the two stories are linked, and that the solution of one applies to the other. Indeed, the problem is basically (in its deep structure) the same. Under this condition, the percentage of study participants who arrive at the correct answer for Dunker's radiation problem rises to 92% (Gick & Holyoak 1980, 1983). Two elements are used: a similar example and an explanation of the common nature of the problem and the solution. Here is how it was done in the Gick and Holyoak (1980) experiment. First, they asked participants to read the general's story as well as two other stories (distractions that had nothing to do with the problem to be solved). Participants had to memorize all three stories and retain key elements of the stories. In this way, the authors ensured that any difficulties in solving the problem posed later were not due to a failure to remember the general's problem. They divided the participants into two groups. One, the experimental group, was instructed to solve the surgeon's problem. They were also given the explicit suggestion to use one of the stories they had learned to help them solve the new problem. The control group did not receive this suggestion. Eleven of the twelve subjects in the experimental group used the "right" solution, corresponding to the general's problem. Ten of these eleven said that the general's example was very helpful in solving the new problem. In the control group, even though they had read and memorized the general's story, three out of fifteen subjects reached the "right" solution - about 20%. The conclusion drawn by the authors is that analogies with distant areas can help in problem solving, but that perceiving and using this analogy is not easy. Being exposed to an analogous example does help, but not as much as being exposed to it while knowing that one must look for a useful analogy in the example in question. In fact, the role of explanation seems to be crucial in this case. Why, in the absence of an explicit suggestion, do subjects tend not to perceive the useful analogy?

4.4.1.2 Similarities and analogies, but at what level?

The concept of the level of abstraction is key in the literature on transfer by analogy. Let us return to the example of the general and the physician. There are several similarities between the two stories. For example, in both cases, the problem has a geographically central position. At a deeper level-actually a more abstract analogy, in both cases it is about defeating an enemy placed in the middle of an environment to be preserved. But if we imagine the two stories without any specific details, ignoring the position of the fortress and that of the tumour, we risk missing a hint of an analogy that could lead us to transfer the solution from one problem to another. But, if we drown in the details of the specific area of each example, we may miss the similarities in this case. For example, we stay with the medical content of one and not the other. It therefore seems reasonable to ask the question of the optimal level of abstraction that facilitates the recognition of similar elements between the two stories - useful for the solution - and therefore makes it possible to use it. How are analogies noticed? How do we manage to connect information, which belongs to distant bodies of knowledge and semantic domains? For a teacher, this type of question is a well-known problem: how do we make a student who has dealt with a knowledge or skill in one subject course understand that the same knowledge or skill can be represented in another course, perhaps in another discipline or on the same subject? The problem of analog transfer, which seems to be crucial for problem solving, is therefore also crucial for academic learning and critical thinking in general. Indeed, a critical thinking skill learned at school is meaningful for the pupil only if he or she is able to mobilize it in the context of his or her daily life: in contact with information, a friend's "snapshot", an observation or an intuition that belongs to him or her.

So we see that the problem of transfer and the problem of analogy are strongly linked[218]. In the literature on analogy, this type of question is translated in terms of *evocation* and *mapping*. Evocation and mapping are considered two fundamental steps in the formulation of an analogy. Evocation is the process during which one story, formulation, text... evokes another. In the example of the general and the surgeon, it is a question of understanding which elements of a story make one think of the other as being significant for the problem to be solved. Mapping then consists of putting similar elements side by side to succeed in seeing how the solution of one can be transformed and translated into a solution for the other. Moreover, evocation and mapping would not be helped by the same type of elements and similarities. Thus, at least according to some of the authors, evocation would rather be helped by superficial similarities, relating to semantic contents. On the other hand, the mapping would rather be supported by similarities in the structure of the two (or more) blocks to be compared. Conversely, the limits of the analogy would be linked to those of the superficial context and to the attraction it exerts on our attention and memory. Caught in the surface elements, we would not be able to perceive what lies deeper. Thus, while the cases of the general and the surgeon are analogous in depth, the surface context is so different that the deep analogy is not perceived because it does not also attract attention.

Gick and Holyoak (1980, 1983) and Holyoak and Koh (1987) thus found that, in this example, the use of the first problem solved to solve the second (the transfer) occurs spontaneously if the two problems have similarities at the surface as well as at the deep structure, that the transfer is not spontaneous but occurs if the two problems differ at the surface and have structural similarities, if the subjects are informed that the first problem is useful for the second. In practice, spontaneous transfer takes place but only under particular conditions; if these do not exist, transfer takes place only if certain less intuitive similarities are made explicit. We will then find these indications from the literature on analogy in the form of practical strategies to promote learning transfer.

In the literature on analogy as well as in the literature on transfer, a distinction is made between deep elements (which belong to the structure of the problem) and surface elements (which should be ignored). We note, for example, that experts in a field are less distracted by the surface of the problem and find it easier to perceive and use the level of the structure. For example, if two problems are superficially similar, but not at the level of structure, experts do not use one to solve the other (Novick 1988). They are at a level of abstraction which is not that of the novice, especially because they possess knowledge formulated in an abstract way (laws, principles, generalizable knowledge).

However, the difficulty remains in clearly defining and distinguishing between surface and deep similarities. In fact, we can each time identify different types of similarities that concern the objects represented in the two problems or stories, the logical relationships with other objects, the actions performed, the events taking place. But the subjects may interpret these properties in different ways due to their prior knowledge and experience, for example (Bassok & Olseth 1995). Therefore, what counts as similarity may be clear to experts, who are at the right level of abstraction when interpreting the problem, but not to novices, who encode situations on the basis of their limited knowledge of the field. Encoding at too specific, too detailed a level can also make it difficult to perceive useful analogy. If the problem lies at the level of encoding - of interpretation - then one strategy for solving problems using the right analogies is to encode the problem correctly - at the right level of abstraction - ignoring the too specific details that belong to the 'surface' of the problem (Gamo, Sander & Richard 2010, Raynal, Clement & Sander 2017).

In conclusion, the use of analogy is a powerful tool for transfer because the former is, in essence, the basis of the latter. However, transfer encounters obstacles specifically related to the ability to use the "right" analogies. Strategies need to be put in place to help perceive and use these analogies. These strategies can be located at the level of formulating problem situations (multiply similarities even on the surface to encourage transfer); they can consist in the explicitation of similarities at the deeper level, useful for resolution; or they can consist in aids to encode the situation at a more abstract and general level.

Let's imagine that we want to help a student recognize certain reasoning biases common to many different situations and contents, such as confirmation bias. The teacher could show him/her examples where there are sufficient surface similarities to recognize the analogy between the examples. Next, the teacher could suggest that the student solve a problem with a previously encountered and solved bias, and explain that the new problem is basically similar to the previous one. Finally, he could help the student to encode each situation presented at a more abstract level, in particular by explaining the reasoning biases, and propose a general formulation of each bias (confirmation bias, selection bias, etc.) in order to help the student recognize them in different situations.

4.4.1.3 Limitations of analogies

While analogy is a powerful means of transferring knowledge and capabilities from old to new content, it has limitations and risks. Analogy can indeed lead in an undesirable direction and evoke knowledge or skills inappropriate to the situation through surface similarities. The case of naïve or preconceived conceptions, which constitute an obstacle to new learning, is well known. For example, many students find it difficult to retain the idea that the differences between the seasons on Earth do not really depend on the distance Earth-Sun but on the inclination of the Earth's axis. Probably part of this resistant misconception is the realization from experience that the distance to a warm body greatly influences the heat felt. An experiential knowledge is thus mobilized to explain a phenomenon which is on an inaccessible scale, by analogy; but the analogy is only superficial and does not take into account significant differences for the phenomenon to be explained. Other experiences show that being exposed to a certain type of problem - and to its solution - can negatively influence the effectiveness of reasoning in problems that are superficially similar but require a different type of solution.

Let's imagine that we have to solve the following problem: we have to measure a certain quantity of eau and we have several jugs of different sizes at our disposal. For example, we need to measure 100 ml of water and we have a decanter that can hold 21 ml, a 127 ml decanter and a 3 ml decanter. What can we do? Then we have to measure 99 ml with 14 ml, 63 ml, 125 ml. Then 5 ml with 18 ml, 43 ml, 10 ml. Then 12 ml with 9 ml, 42 ml, 6 ml. Then 31 ml with 20 ml, 59 ml, 4 ml. Then 20 ml with 23 ml, 49 ml, 3 ml. Then 18 ml with 15 ml, 39 ml, 3 ml. Finally, we have to measure 25 ml with 28 ml, 76 ml and 3 ml jugs. People who solve all the problems in series usually have difficulty with the last one. However, the last one is not a problem if it is presented in isolation. This means that the solutions given beforehand influence our ability to look for a solution in a new direction and may prevent us from finding the right solution if we are anchored by analogy on an inappropriate solution (Luchins & Luchins 1970).

This problem with analogies suggests that they need to be properly checked. Not all of them are productive. The problem of negative transfer also alerts us to the importance of properly evaluating the strategies that students use in problem solving. When they stumble or fall back on the wrong solution, the cause may be the use of the wrong analogy, called for by the circumstances or particularly present in the student's mind at the time.

4.4.2 Benefitss of contexts variation

The context acts on several levels. There are many examples of "street learning" that is not transferred to more academic, classroom situations and, conversely, classroom learning that is not re-exploited in everyday life. For example, children may be able to perform certain calculations in the context of selling objects on the street without being able to respond to similar problems in the classroom (Carraher 1986; Bransford et al. 2000). The fact that context seems to have a negative impact on transfer suggests that concrete learning, rooted in real life cases, is potentially negative. Indeed, learning that is rooted in real problems to be solved is more resistant to transfer to new situations. However, these seem at the same time to facilitate understanding and learning - for example, in fields such as medicine, learning from concrete cases and problem solving seems to be an effective method. How can this apparent conflict between understanding and reuse be resolved? One way, also basically present in the case of the general and the surgeon, is to multiply the contexts - starting from a specific case, then proposing similar cases - and ask to compare these contexts. Another is to ask to produce similar cases. Minervino, Olguin and Trench (2017) have proposed the case of the general and other similar cases to their subjects and asked one of their groups to produce a similar case. They then have placed all the groups before the surgeon's problem. They have shown that production positively influences the ability to remember deep analogies over time. If the transfer to daily life is the one aimed at, it is also necessary to include, among the contexts presented, contexts of daily life to which we hope to see the transfer take place[219]. This includes contents towards which we want the transfer to take place, but also ways of doing and working - for example, knowing how to collaborate with others.

4.4.3 Benefits of explanation

The case of the general and the surgeon also indicates that varying contexts is not sufficient or necessarily effective if the learner is not able to perceive that there are analogies between these contexts at a deep level. It is therefore a question of pointing out, highlighting and making these analogies perceptible. This can be done through different forms of explanation. For example, in the case of the study conducted by Gick and Holyoak (1981, 1983), we could point out that there is a common structure or make it even more explicit by formulating the general principle underlying the two cases. Thus by using direct, explicit teaching of the general concept to be retained. A historical example illustrates this second way of doing things and its positive effects. Children have to centre a target with darts; the target is covered by a few centimetres of water, and its visual position is therefore displaced by refraction. One group of children receives a short lesson on refraction; the other does not. Both practice on the same target and are equally successful. But, when the amount of water changes, the group that has received an explicit explanation about the principle of refraction is more successful than the other (Scholkow & Judd 1908, in Bransford et al. 2000). In this case, therefore, it is a matter of mixing case-based learning with instruction that leads to a more abstract representation of the problem (Singley & Anderson 1989; Bransford et al 2000).

In some cases, the explicit instruction is functional because it makes it possible to isolate a discriminating criterion that applies to the greatest number of cases. A well-known case is that of distinguishing the sex of young chicks. This perceptive task is particularly difficult for novices, but improves greatly with expertise: experts are able to distinguish the sex, male or female, of a thousand individuals in one hour with 98% accuracy. However, as the cloaca of newborns is similar for males and females, novices are not sure what they need to look for to distinguish between the two: a critical region exists, which is rounder in males. Biederman & Shiffrar (1987) show that, if we clearly explain this difference to naïve subjects, their performance improves to the point where they immediately resemble that of experts, compared to simply re-training them to observe without being instructed on what to look for. Providing criteria that the experts use (implicitly or explicitly) therefore allows them to improve more quickly and to connect several examples together. In general, teaching learners to represent a problem at a more abstract level allows the different cases and concrete examples presented to cease to exist as individual cases and enter into a larger family or scheme (Gick and Holyoak 1983; Bransford et al. 2000[220]).

4.4.4 Low and high

The variation of contexts, aids or supports for abstraction and the role of similarities are very present in one of the classical approaches to transfer, that of (Salomon & Perkins 1987). They distinguished two paths or modalities conducive to transfer.

On the one hand, the low track goes through extensive practice and, in a wide variety of situations, up to automaticity. Generalization is produced through this variety of situations in which the ability is practiced. However, the transfer is highly dependent on the presence of surface, perceptual similarities. If these are not strong enough, they must be called and made evident. Salomon and Perkins offer the example of procedural learning par excellence: that of driving. Having learned to drive a car and having done so under a variety of conditions makes it easier to learn to drive a truck: certain skills can be transferred. The new context (that of driving the truck) is indeed so perceptually similar to the original learning context that it almost automatically calls and triggers the behavioural patterns and routines learned in the context of the car.

On the other hand, the high way is through a voluntary and reflexive exercise. It is a question of disregarding the context in order to extract a general principle from it. The fundamental element in this case is the abstraction with explicitation of a principle - the de-contextualization of the principle to be transferred. This can be used in two ways. On the one hand, the learner can extract an abstract principle and keep it in mind in preparation for future situations in which to apply it; for example, he can learn a principle of physics and ask himself in what other contexts it applies. On the other hand, the learner may find him/herself in a new situation with a problem to solve, extract regularities and characteristics from it, and then look for correspondences with similar situations already encountered. In both cases, it is a matter of ignoring the context and looking for connections with other contexts and contents. The high road therefore does not depend on the existence of superficial similarities - at the limit these can become confusing - but on the search for analogies in the deep structure of the problems - which is not always easy.

The two modes can be combined. It is possible to practise in a capacity until it is automated (a probability calculation for example), then to think about the principles and usefulness of this type of calculation to be applied in different situations. However, distinguishing the two modes is useful to know in which situations transfer is possible or difficult. When it goes through the automatic, low channel, transfer does not pose major problems. In any case, it is probably a close transfer - because of the similarity of the situations. Transfer through the high channel allows it to link more distant contexts, so it opens the way to distant transfer, but requires an effort of will to abstract and the ability to create connections. The teacher can help the student to transfer both low and high.

Firstly, it is about making available to the learner the situations we want to connect. Let us imagine wanting the student to be able to one day reuse the knowledge learned in history to analyse contemporary political facts, or those learned in biology to an ecological context. One possibility to help him/her to achieve this transfer is to present contemporary and ecological facts in the same context in which historical facts and knowledge in biology are presented. This involves creating learning situations that contain or are similar to the desired one.

Secondly, it is a matter of helping the pupil to build links between these situations, without waiting for him to do it himself and alone: for example, by providing him with general principles and explaining how they are relevant to the context, or by helping him to seek these generalisations himself (by asking questions, soliciting analogies, inviting him to apply the same principle to different areas, to different subjects). In other words, the aim is to facilitate abstraction.

In English, these two techniques are called "*hugging*" and "*bridging*". The aim is for the teacher to use them in a systematic way, knowing when and why they are both put in place, to facilitate what kind of learning and transfer - close or far, low or high way. Salomon and Perkins (1987) suggest that teachers should not keep these techniques to themselves; that is, use them to support student learning and transfer, and share them with the students themselves. Convince students that they can ask themselves how they can learn to transfer their learning and adopt techniques to help themselves transfer in the future: get used to thinking about similarities, for example, between what they do in one subject area and another; between what they do at school and in daily life. To achieve this, however, it seems necessary to think of all education in the light of transfer. For example, to give more importance to transdisciplinary or interdisciplinary notions that create the context (by way of bottom-up thinking) to transfer, or to emphasize basic principles of reasoning and its limits that are inevitably common to all subjects and disciplines: natural tendencies such as getting stuck in one's opinions, not taking alternative hypotheses into consideration, etc., are not always taken into account. Making these tendencies explicit, ignoring the context and linking them to various concrete cases helps to develop the conditions for a high-level transfer, by high route, of skills general enough to invest in any type of learning. Other principles and rules lie halfway between similarities in content and metacognitive abstractions: the need for measures and evidence, rules such as those of method, which are easily applicable to several fields, including but not limited to science[221].

4.4.5 A lot of practice, but not just any practice

Tim Van Gelder - a philosopher and cognitive scientist interested in CT - compares critical thinking to ballet (Van Gelder 2005 [222]). He argues that critical thinking is not an evolutionarily selected ability or set of abilities because it is not necessary for survival in an environment such as that in which our ancestors evolved. Rather, we would be organisms that seek patterns and coherent narratives, without questioning whether the pattern is real and the story is true. We can certainly discuss whether there are any critical thinking skills that do not have an ancient evolutionary origin, such as those that lead us to look for patterns and listen to stories - because not all of these patterns and stories actually have the same appeal to us. However, we can accept as a first approximation this simplified framework, which leads Van Gelder to propose that the critical mind must be voluntarily, and with effort, learned, just as we learn ballet, tennis or a second language. Indeed, Van Gelder explains that, like these abilities, even though the critical mind is composed of more elementary components, learning is necessary to combine them correctly - and not only to acquire them individually. For this reason, learning to think critically takes time and dedicated practice. It is therefore inspired by the field of expertise, in particular the approach developed by Ericsson and called "deliberate practice" (Ericsson, Krampe & Tesch-Römer 1993). According to this approach, anyone can develop his or her performance in a field and become an expert in it, through a long (5,000 to 10,000 hours in total, 10 years for 4 hours of daily practice) practice which also has these characteristics: it is carried out with concentration and with the aim of improving; it uses specific exercises to improve performance; its exercises become progressively more complicated in order to reach a higher level; it is guided and evaluated, crowned by feedback[223].

Van Gelder acknowledges that transfer remains a crucial problem for critical thinking - both because without transfer, this learning loses its meaning and because the transfer of this set of abilities appears particularly difficult and uncertain[224]. So the first step is to recognize the problem, the second is to teach transfer without expecting it to happen by itself.

4.4.6 Metacognition and metacognitive learning

Transfer can be described as the ability to apply knowledge to a new field or context in relation to that of learning. Within this definition, transfer is a discrete phenomenon, which may or may not take place. We have seen, however, that it is not always easy without help - especially without the suggestion to look in a certain direction and look for certain similarities. This indicates that transfer can be a matter of degrees. We may, for example, need more or less help to transfer our knowledge to a new area (generic or specific, limited or numerous aids, etc.). We can also consider transfer as the ability to learn in a new field, based on learning that has been done elsewhere and beforehand (Bransford et al. 2000). These considerations translate into ways of assessing transfer, for example, through measures based on the number and type of 'pushes' that are required to achieve transfer.

It would appear, however, that adding metacognitive approaches to instruction helps transfer, even in the absence of "nudge". These metacognitive approaches vary greatly in format, but they share the idea of focusing the learner's attention on the strategies he or she adopts during learning, the progress he or she makes, and the resources he or she mobilizes (Bransford et al. 2000). Examples of metacognitive approaches are actually examples of learning by teaching or reciprocal teaching, where students acquire knowledge in order to pass it on to others. In the process, learners are led to reflect on what facilitates learning and understanding and to evaluate their strategies (Brown, Palincsar & Armbruster 1984).

A consideration of the effectiveness of teaching methods should be added to these elements. The practice of metacognition in the classroom (reflective practice on what has been learned, explicitness, assessment of learning, learning metacognitive strategies for better learning) has recently received a great deal of attention, particularly following the positive evaluations by John Hattie (Hattie, Biggs & Purdie 1996, Hattie & Donoghue 2016; https://visible-learning.org/hattie-ranking-influences-effect-sizes-learning-achievement/) and the *Education Endowment Fund (EEF*,https://educationendowmentfoundation.org.uk/tools/guidance-reports/metacognition-and-self-regulated-learning/). Like explicit instruction, the teaching of learning strategies and metacognitive strategies is an important - even dominant - part of the current educational landscape. We can point out that the EEF report on metacognitive strategies in the classroom considers metacognitive strategies to be strongly connected to learning content[225]. Introducing explicit, reflexive, self-assessment practices could also help the CE: how do we know? What evidence do we rely on? How well do we know this area, to the extent that we feel confident? Are we right to be so assertive? Is this a situation with risk factors for being wrong (emotional investment, lack of knowledge, etc.)?

4.4.7 Argumentation as a support for CT and as a pedagogical tool

The notion of argumentation (or good argumentation) is often associated with that of EC. For some authors, it is important to adopt argumentative methods to improve argumentation skills but also thinking skills, especially from the age of middle school.

4.4.7.1 The argumentative approach (1)

Deanna Kuhn argues that inserting argumentation into the school curriculum is as fundamental as a mathematics or history course[226.] She sees argumentation as a general ability which, in turn, includes different forms of reasoning. The non-disciplinary nature of argumentation, unrelated to any specific content, makes it more difficult to teach, but Kuhn does not see this as an objection. Indeed, she has developed a method of teaching argumentation that she considers effective for students of middle school age. The method, which is based on training in debate, argument analysis and production, would, according to Kuhn, have a positive effect on written argumentation at the individual level, on the ability of a group to lead debates as well as, what interests us here, on the ability to analyse arguments and supporting evidence and thus to evaluate a thesis (Kuhn, Hemberger & Khait 2015). The proposed approach is of the implicit type; the method is therefore based on the exercise of dialogue and not on explicit instruction (Kuhn, Hemberger & Khait 2015). However, the transcripts of the dialogues are used for debriefing, and reflective activities are proposed in order to analyse how the debates took place and what needs to be learned from them.

The method does not take into account the truth of an argument, but rather rhetorical abilities and openness to the positions of others: the student must initially learn to pay attention to the positions of others and seek to respond with counter-arguments, the aim being to weaken the opponent's positions. Only once this is acquired, the student learns to use evidence to strengthen or weaken positions. In fact, the student could learn to strengthen his arguments by selecting supporting evidence and ignoring evidence against his position. The opponent must do the counter-argument work. From the point of view of educating the critical mind, this

would not be an individual objective, but only for a group of people who are in dialogue or even debating from opposite positions (Kuhn, Hemberger & Khait 2015[227]). Pupils debate among themselves using software that presents questions ranging from everyday questions to ethical questions related to political or health decisions. They work in pairs to develop a common argument "against" two other students. The whole intervention is based on phases of argumentative debate in small groups (with search for arguments and counter-arguments), followed by a debate in large groups and then a debriefing. The cycle starts again at the end of each intervention (thirteen weeks at the rate of two meetings per week). Some of the classes that took part in the experiment worked on the project for two years, others for three years. The project is led by an external facilitator. Control classes follow a programme of equal intensity but rather based on Lipman's approach to philosophy for children - without structured dialogue, small group work and preparation of counter-arguments - and are supervised by a teacher from the school. The impact measurement is qualitative, micro-genetic, based on the analysis of the productions.
Beyond the difficulty of commenting on the results obtained by this program, which is cumbersome to implement, and although the objective is to improve the ability to discuss based on solid arguments, this type of intervention does not seem to us to meet the definition of CT proposed in this document. Indeed, students choose their arguments based on their knowledge, which is probably partial. Moreover, they are not led to read before debating in order to become informed (to avoid restricting the debate or providing information that would not immediately appear relevant and interesting). Students identify themselves the questions that require more information, and the teacher provides them in a more advanced session with the information that he or she selects in response to these questions. At no time are students confronted with the need to assess the quality of the information, its weight in terms of supporting evidence and the reliability of the sources (Kuhn, Hemberger & Khait 2015[228]).

Kuhn has also proposed other methodological approaches to argumentation, which she relates directly to CT. For example, she has explored the impact of argumentation on the CT abilities of junior high school students (in fact, the impact of a group of explicit lessons on the analysis of arguments, Kuhn 2015). The course is considered a philosophy course and consists of eight thirty-minute sessions, at the rate of two meetings per week, during which a speaker explicitly explains the value of listening to the arguments of others (open-mindedness), how to construct counter-arguments, the importance of questioning oneself about one's beliefs, and the importance of analyzing the strength of arguments on the basis of their construction and supporting evidence. The sessions also include moments of debate and discussion among students with opposing views in order to collaboratively analyze each other's counter-arguments. The impact measure consists of a pre-test/post-test consisting of six open-ended questions regarding the analysis of arguments. The questions are proposed in the form of an interview. Students are asked to justify their opinion on an option that does not require special knowledge to decide (e.g. the length of the school day). Evaluators assess (qualitatively) the student's ability to give an opinion, to provide supporting reasons or to explain how he or she would go about seeking supporting evidence (what kind of research would be needed, what sources could be interviewed, such as establishing their reliability, etc.), to consider the opposing options and the arguments in favour of them. Other questions assess students' metacognition, thinking, and the meaning they give to the idea of "thinking better".

In only one school did 24 student volunteers participate in the study (26 minus two who did not answer the questions). They were divided into three groups, two treatment and one control group. The study presents a positive outcome of the philosophy course, showing an evolution of the quality of answers between pre-test and post-test only for the students in the treatment groups. However, this result is actually of low evidential value given the poor number of participants and the qualitative nature of the measure. No measurement is provided over time, nor in an "ecological" setting different from that in which the students were trained in class during the seminar. In particular, the format of the test corresponds to that of the last day of the seminar, when the students are led to debate among themselves about the existence of Santa Claus.

4.4.7.2 The argumentative approach (2)

Hugo Mercier also offers an argumentative approach to reasoning and solving intellectual tasks. His approach is based on a vision of reason as argumentative by its nature: the capacities relating to "reason" (the search for and presentation of arguments in support of a position) and reasoning (the search for solutions to an intellectual problem, thanks to available knowledge) would have developed in response to a social need: that of putting forward one's arguments, of persuading. When we look for reasons, we do not seek to make the best decision, but to convince others that we are correct. a. our individual and solitary reasoning would often mislead us (Mercier & Sperber 2011; Mercier 2016; Boku, Yama & Mercier 2018); b. the[229]group situation, argumentative, would give better results in terms of solving intellectual tasks (Boku, Yama & Mercier 2018).

As for performance transfer, the literature on argumentation is not clear. It would show that participating in an argumentative discussion situation increases the chances of solving similar problems in the future (immediately afterwards, in the post-test phase that follows). In the case of the experiments described by Mercier and his colleagues (e.g. 230Trouche, Sander & Mercier 2014; Boku, Yama & Mercier 2018), it is a question of close transfer: the transfer takes place on the same structure of the problem, with a different content. However, the transfer is not in itself part of the predictions of the theory, because the experiments are not supposed to test an educational approach but a situation of solving a reasoning problem. Indeed, close transfer allows to verify that the participants have understood why the right solution to the problem is the one given at the end of the debate, but not whether they know how to use different reasoning strategies in different problems thanks to having taken part in the experiment. Participation in debate or argumentative discussion activities may therefore only serve to achieve better performance in the specific case, but not to learn a transferable skill (Boku Yama & Mercier 2018). The tests do not include a control group that would have the same time as the treatment group to solve the problem (participants try to solve the problem individually for five minutes, and this serves as a control group for the 10-minute group phase, after which participants still have five minutes to answer the riddle) (Boku, Yama & Mercier 2018).

4.4.7.2.1 Argumentation and evaluation of arguments

The use of argumentation in the range of methods for teaching CT is, however, to be taken seriously for the following reason, developed by Mercier (2016): the evaluation of arguments in support of a position seems to be more favoured in a discussion situation than in an individual situation. In an individual situation, it would be better to be "lazy" in confirming one's own positions[231]. On the other hand, in a social situation, the individual would be inclined not only to look for good arguments in favour of his position but also to evaluate those of the opponent in a more objective and demanding way than he would do with himself. Asking someone to evaluate their own arguments would therefore be less effective than asking them to evaluate the arguments of others. (It is, however, more difficult, in an argumentative framework, to measure the ability to evaluate arguments than the ability to produce good arguments[232]). Furthermore, Mercier (2016) defends the idea that, in everyday life, an average subject is relatively effective in evaluating the strength of arguments brought by others. In particular, Mercier (2016) cites studies that show that interlocutors identify weak arguments because they are based on argumentative misconceptions (an argument of authority relating to a non-expert and not necessarily benevolent source) or on logical errors[233]. The situation of dialogue specific to argumentative exchanges would therefore make it possible to make better use of these capacities compared to a condition of solitary evaluation. It would then represent a particularly favourable condition for the exercise - if not the education - of the CT.

Cooperative learning appears to have positive results in education and in the development of young children's cognitive abilities (Slavin 2014; Sills, Rowse & Emerson 2016). Kyndt and colleagues (2013) have conducted a meta-analysis of studies on cooperative learning since 1995 and confirmed positive effects on learning and attitudes.

4.5. Syntesis and practical consequences for CT education

Contemporary psychology seems sceptical about the possibility of learning heuristics or general rules and applying them independently of content. Even the most optimistic (such as Richard Nisbett) recognize the limits of this type of generalization: the need to already have an intuitive system of rules to be substituted with a more sophisticated one, transparency of the problem (which lets perceive the type of rules to be applied), limitation of systems of general rules to few typologies (statistical rules, causality, certain social rules). However, it does not seem irrational to maintain the **hope of developing certain cognitive skills in a general way, beyond a given content and context, or at least to see them transferred in several different situations**, as long as these skills:

- are based on strategies that we can learn and then practice consistently in different contexts;

- are expressed in an abstract way so as to find them more easily and develop shortcuts in our memory.

This hope must take into account the different forms of transfer, which may be more or less remote, because of the similarity between the situations involved. We must not ignore the role played by knowledge both in the ability to solve problems or think effectively, and in the constraints that this places on transfer in new situations where we lack the necessary knowledge.

The learning of general thinking skills, such as CT, does not seem to be divorced from the acquisition of knowledge. To think well in a new field, one also needs to acquire domain knowledge of sufficient quality and quantity. Having a rich and varied knowledge base is a necessary condition for good thinking. General heuristics are not sufficient.

We can hope to transfer - at least to some degree - the skills acquired in practice from one particular field to another, if this acquisition has followed learning principles for transfer. Indeed, transfer does not happen by itself or automatically. Where it resembles an automatic transfer, two conditions are in fact probably present: on the one hand, extensive practice has allowed the automation of the ability in one context; on the other hand, the new context has strong similarities with the context of acquisition, perceptual attractors that strongly facilitate the transfer and, we might say, even call it "against" will.
In other cases of transfer, general heuristics or abstract principles have been voluntarily produced and hung up on concrete contexts, on practices. They seem necessary to facilitate, or even make possible, the transfer. But we cannot expect abstractions and knowledge bases or concrete situations of application to come together on their own, without the help of appropriate forms of learning and possibly without punctual and local aids inviting to be taken into account in relation to each other (suggestions, external aids, facilitations in the form of similar contexts even on the surface).

Education for CT is the story of a third way: education does not need to give up general abilities, to retreat only into the acquisition of specific and local knowledge for fear of

"wasting time" unnecessarily; nor does it have to choose to develop thinking skills while giving up providing students with a rich knowledge base; the principles of reasoning, critical thinking and methods of investigation cannot be taught and learned in an abstract way, nor should they be practised without being accompanied by general principles and heuristics and by reflection on how to transfer these heuristics to a new context.

In everyday life, we are often faced with the following situation: we cannot expect to acquire in-depth knowledge in all the areas we encounter, but we are led to take a position, express an opinion and decide in one direction or another. It then seems necessary to adopt a principle of reality. It is unlikely, if not impossible, to imagine that we could also be able to think "expertly" in any field. **The aim is not perfection, but an improvement in thinking skills and their more satisfactory application in contexts and in the face of content of personal or general interest.**

Let us return to the notion of weak method and weak results. By using general heuristic principles in a new context, we can obtain a transfer and succeed in giving a sufficiently satisfactory answer - compared to someone who would not have these heuristics. However, the result will probably not be as satisfactory as if, in addition to general principles, we could also rely on domain knowledge. It will then be a compromise between the cost of acquiring, in addition to general principles, new domain knowledge and the gain of an optimal response to the new situation.

4.5.1 Practical tips to promote the transfer of CT

Practical advice for teaching CT emerges from this synthesis. However, we must stress that they are based on qualitative elements of literature review and therefore need to be tested and tested with appropriate methods to measure their effectiveness:

1. *Teach critical thinking skills and strategies/knowledge explicitly and in an assumed way.*
    - Teach critical thinking skills in an assertive way: don't expect that by infusing them into "curricula" they will be understood, generalized and transferred.
2. *Teach for transfer.*

- Teach general principles and general strategies in relation to concrete cases to show how they can be put into practice.
- Repeatedly and variably practice a capacity up to automation.
    - Rehearse in varied but close contexts.
        - Multiple disciplinary examples.
    - Rehearse in contexts where we want the transfer to take place and highlight the similarities.
        - Examples from everyday life.
- Make abstract and explicit to encourage reflexive transfer and possibly help with suggestions, questions, etc. if this does not come alone.
- To make people think about what we are doing by mobilizing metacognition.
- Use analogy at the right level of depth and relevance.

3. *Rely on a critical "social" mind.*
    - Use argumentative debate situations, group activities.
4. *Rely on external scaffolds.*
    - Use tools to learn how to transfer but also to learn how to learn how to transfer: make a list of those we have used in different situations.
    - Explain the reasons for using this type of tool and encourage people to make it their own.

5. Pedagogical advice: do's and don'ts in CT education

The theoretical contributions we have mobilized in this report allow us to identify a set of pedagogical opinions concerning CT education: some paths to be tried and others that are not theoretical or practical, or both, lead to rejection. We shall begin with the latter.

5.1 Don'ts

5.1.1 Don't reduce CT education to information and media literacy

A paper published in 2012 showed a correlation between the number of Nobel prizes obtained in a country and the annual chocolate consumption of the population of that country (Messerli, 2012; see also Maurage et al. 2013). This type of graph prompts the reader to make an almost immediate inference, shifting from the finding of a correlation to a causal link: we imagine, even for a few moments, that chocolate consumption is the cause of obtaining Nobel Prizes, perhaps because chocolate contains molecules with properties that affect the development of our intellectual capacities. This article has been taken up in various French and foreign media: the headlines are sometimes cautious in their wording ("The more a country eats chocolate, the more Nobel Prizes it has"), others less cautious ("Eat chocolate to get the Nobel Prize"). Such an anecdote illustrates the fact that we have no need of the media to come to these kinds of conclusions. The media plays an undeniable role in circulating this information, but the information itself is seductive in the sense that it has characteristics that make it spread quickly. The first of these characteristics is cited above: we infer a causal relationship on the basis of indices whose correlation can lead us to illusions of causality (Matute et al. 2015). This tendency is somehow prevalent in the existence of this "false information".

In the third part of this report, we stressed the importance of equipping students with specific criteria and knowledge adapted to the context, which allow for a more refined evaluation of information sources (in relation to the criteria and indices used spontaneously). An education of the critical mind that would only focus on media education would however be insufficient: in the example presented above, the information was initially proposed by a doctor in a prestigious journal. Focusing on the source (solid in this case) or on the advanced evidence (having enough knowledge in biology to discuss the facts underlying this thesis is delicate... the students might even manage to convince themselves that the thesis is credible) would prove insufficient, because the heart of the problem lies elsewhere. Moreover, without taking more precautions, students might think that it is necessary to start doubting every piece of information, including that provided in scientific articles. In reality, the author of the article does not conclude that there is a causal link, as has been reported in some media.

The conclusion of this reflection is not that we should deny the interest of media education. On the contrary, students must be made aware of the characteristics of the new media. However, this learning must be linked to a reflection on what, in our own cognitive functioning, can lead us into error. For example, we all spontaneously look for causes of the phenomena we observe, and finding correlations is a good clue to do so. But finding a correlation is not enough: apart from other clues, it is still a weak level of proof. How could we be sure of a cause and effect relationship? An experimental test or the search for causal mechanisms would provide stronger evidence. Rather than simply constructing a dark picture of the media world in the minds of students, the goal is to give them tools to evaluate the information they receive.

5.1.2. Be careful not to give the impression that what matters in an argumentative debate is to present one's arguments well

In the fourth part of this report, we stressed the importance of mobilizing learning situations in which students are led to debate and argue. Collective reasoning could encourage their ability to identify bad arguments, or bad argumentative structures, whereas the same approach conducted alone would lead them to withdraw into their positions. It is important to note, however, that argumentation work alone is not sufficient to educate CT.

One aspect to consider when working on the argument is the risk of an unframed recourse to debate. Some teachers might decide to organize classroom debates on lively social issues (for example, the danger of vaccines or GMOs) to convey general messages about attitudes: some teachers might take the opportunity to show that it is easy to lie, to develop false arguments, and thus invite students to 'beware'; in the opposite way, others might state that there is not one 'one truth' but a set of points of view. The teacher would then invite students to be open-minded, not to remain entrenched in their positions, and finally to be open to contrary ideas on the pretext that no knowledge is fixed (especially not in the field of science). What could we blame for such messages?

Firstly, such general advice is difficult to put into practice and to apply in specific contexts (beware of whom?). To be demanding in what way? How can we be open-minded?).

Second, such advice can lead to misinterpretation. To close to all information blindly or, on the contrary, to consider that all knowledge can be questioned is to overlook the fact that not all opinions are equal. Because no thesis is completely certain, we group together in the same category rigorously established knowledge, verified by bundles of evidence, and unsupported personal points of view. By asking students to take turns defending "anti" and "pro-vaccine" positions, for example, the teacher would encourage them to exercise their argumentative skills by denying the superiority of some arguments (those supported by facts) over others. Remaining open to all positions can lead to a posture of doubt, a relativism that is opposed to critical thinking. On the other hand, to be suspicious by default, by over-vigilance, is also to refuse to acknowledge the superiority of certain information. This posture invites one to withdraw into one's own personal convictions, without discernment.

On the contrary, teaching debate should create the conditions for students to understand that not all points of view on a subject are the same: not all are acceptable, nor do all deserve our suspicion. The relative epistemic value of the information at our disposal is what makes it possible to decide. Debate should lead students to decide between points of view on the basis of facts, rather than simply practicing defending a position at all costs.

Finally, it is important to realize that, in a debate, we often associate positions of preference with considerations of fact, supported by evidence. Whether or not to accept that society should use a particular technology from an ethical perspective is a matter of preference, of personal positions. These preferences are *a priori* valid and are debatable. The safety of a given vaccine, on the other hand, is a scientific fact and is not debated in terms of personal preferences, but in terms of the evidence underlying the assertion. Placing students in a debating position is prepared and accompanied, not to bring students into a free and open exchange on all topics, but to teach them to distinguish between preferences, opinions, knowledge, facts and different levels of evidence.

5.1.3. Do not limit yourself to hot topics

We might be tempted to focus CT teaching around major scientific theories such as the theory of evolution or global warming. Such a choice may be counterproductive.

First of all, in addition to cognitive engagement, these discussion themes include an aspect of belonging to a certain social group. It becomes particularly difficult to divest oneself of one's ideas in such a context (see the second part of this report). Thus, a student who adopts a position that runs counter to a scientific theory may do so on a completely different basis than that of reasoning, arguments, and facts. In this case, inexorably bringing the discussion back to the facts or the construction of scientific theories would not necessarily be sufficient, as other elements would then be at stake.

Secondly, there is a risk in seeking to undo a "myth", a false belief: that of reinforcing this position. Some students may not be familiar with a pseudo-thesis (such as that of *intelligent design*) and, in this case, teaching that aims to deconstruct the pseudo-thesis would first expose students to this content. It is also possible that students exposed in this way would forget, after some time, that the teacher had specified that this was false information, and would only remember the content. More generally, trying to deconstruct false theories is a risky exercise that can produce the opposite effect to that hoped for (Cook, 2017).

Finally, designing the teaching of critical thinking in the form of a one-off project or on a sensitive theme may focus the students' attention on the context in question (see Part Four of this report). In such a situation, the tools presented by the teacher are then very context-dependent or so associated with the context that students have great difficulty in transferring it to other scientific theories, and even more so to situations of everyday life. It is therefore not possible to consider that the teacher has developed the students' CT. The teaching of CT cannot therefore be based entirely on teaching around lively questions.

Does this mean that themes such as the theory of evolution should not be discussed? No, of course it doesn't. When tackling more delicate subjects, the teacher can remobilize all the tools of scientific and critical thinking that he or she has taught up to now. It would be an illusion to think that students could easily import them and ignore all the obstacles to their understanding of the theory of evolution. There are several reasons for this:

- when students discover this theme, they have little solid knowledge. It is impossible to exercise one's critical thinking outside of one's knowledge.
- The highly charged context of this theme makes it particularly difficult to identify the deep structure of the obstacles encountered (for example, the role of chance, reasoning by consequences, etc.).
- Each theory has its own pitfalls, its own obstacles, which are difficult to find elsewhere and which hinder global understanding as long as they persist.

Nevertheless, it is reasonable to think that the upstream use of transferable tools (both knowledge of our own limits and more epistemological knowledge of how science works, of the nature of scientific fact in relation to opinion, etc.) can favourably influence adherence to major scientific theories.

5.1.4. Do not limit CT instruction to an awareness of cognitive biases, and avoid lists of biases

A large body of general public (as well as scientific) literature on CT focuses on cognitive biases and errors in judgement that would allow humans to have only limited rationality (see Part 3 of this report). Thus, long lists of biases that highlight how poorly we reason are popularized (see, for example, Wikipedia. *List of cognitive biases*).

In the third part of this report, we have shown the value of understanding that certain circumstances put us at greater risk of error than others, and the value of knowing how to identify cognitive barriers to a fair assessment of the quality of the information available and to calibrating one's confidence in this information.

However, there are limits to such a pedagogical approach. First of all, it may convey the idea that our cognition is actually of little use to us. Such global considerations, such as the idea that "all information on the Internet is wrong", do not provide a fair discourse on which to build. Yet we recognize many faces, many emotions, we know facts that we can attest to with our senses (the apple falls well to the ground, in accordance with the theory of gravity). In short, our natural tools serve us well most of the time. It is true, however, that some scientific theories are counter-intuitive, that our approach to statistics and probabilities is limited, and that our reasoning is based on heuristics that are not always appropriate. The pedagogical effort to be made is therefore not to give the student a pessimistic list of all the categories of errors we would make at any given moment, but rather to point out that certain disciplinary fields or certain situations (such as a graph showing a correlation) make us misjudge the quality of information and that it is important to learn to better identify and therefore anticipate these situations, or at least to lower confidence in one's opinions when circumstances are more conducive to assessment errors.

Let us stress that it is not a question - in the context of a CT education - of leading the student to systematically lower his confidence, at the risk of simply sowing doubt in his mind. Rather, it is a matter of getting the student to recognize situations in which he or she can legitimately trust and those in which the information is less reliable or even frankly suspect. Armed with this knowledge, the student can thus increase his or her vigilance (or lower his or her confidence) in these very specific situations, and not blindly, at any time.

5.1.5. Don’t underestimate the difficulties of transfer

We would hope that teaching the scientific method is sufficient in itself to enable students to recognize information that is supported by solid evidence and distinguish it from anecdotes. Students would discover the tools of science in the classroom and convert them into tools to guide their choices of the best information. Unfortunately, we have outlined all the reasons why we believe that such spontaneous transfer is not achieved (see Part 4). Several authors point out the difficulty for the learner to abstract from the context in which he or she discovered a notion, in order to remobilize it in a new situation. Instead, the learner tends to remain blind to the deep structures of problems and caught up in their superficial structures. The challenge of transfer is all the more delicate as the distance between the domains (school domain and daily life) is great.

A specific effort dedicated to transfer must therefore accompany any science (or other discipline) course that presents an information assessment tool. Especially since, as we have pointed out, the exercise of critical thinking is also dependent on knowledge. CT education could therefore not be at the expense of existing disciplinary courses.

It is therefore not a question of minimizing the teaching of fundamental contents, but of finding a place for explicit discourse on more abstract aspects. Thus, whenever he or she is in a position to do so, the teacher should stress that two assertions, opinions or theses are not equal but, on the contrary, that one is better than the other for a given reason. This reason may be centred on the facts (one thesis is supported by facts, the other is not), or on the methods that produced the evidence to support each thesis (for example, one relies on a body of observations, the other on a few anecdotal accounts), or the sources that relay them. The teacher therefore takes advantage of opportunities within his or her classical teaching to teach students strategies for better thinking and better evaluation of information (content and sources).

This teaching is evolutionary. For example, assessing the expertise of a source is not done in the same way at seven and fifteen years of age. Initially, the teacher seeks to awaken students to the idea that clues can guide us more or less effectively in assessing the source's expertise. Later on, he or she can talk about the way science works, the different types of expertise, etc.

Even with all this, the difficulty of the transfer should not be underestimated. While it is trivial to consider that several observations are better than one, or that an expert source is better than a non-expert source, practicing the presence or absence of such criteria in everyday situations is a real challenge. Such attitudes remain costly to put in place in real-life contexts where we lose sight of their relevance and often seek to move quickly. Only regular and explicit learning of the same considerations in a very large number of situations would allow us to hope for an effective transfer in complex situations of daily life. How should it be organized?

In primary school as well as in middle and high school, one solution is represented by the explicit connections between disciplines. For example, the teacher of a discipline conducts his or her course in a classical manner, and then, at the end of the course, sets aside time for more general considerations of evidence or sources in support of a thesis stated in the course. He uses a first recourse favouring transfer: explicitation. He draws the students' attention a first time, through a general discourse, to the deep structure of the problem, beyond the context in which it was encountered. A little later, the teacher of another subject (or the same teacher, of course, in the context of primary education) in turn carries out a course on a very specific content, different from the first. In turn, he or she also devotes a few minutes to present a more general and explicit message. He also shows the link with what the students discovered

a few days earlier, encouraging the transfer to another context. The students thus receive in turn a learning process based on concrete cases and a discourse that helps them visualize the problem at a deeper level. Teachers can go further by proposing to pupils situations (particularly in the form of formative evaluation such as quizzes) from everyday life that mobilize the same obstacles and the same solutions. Pupils can also be encouraged to find such examples themselves. This stage is critical: not only does it allow for the work of the competence to be targeted at the area of the end (everyday life), but it also helps students to understand what the teacher expects of them and to focus their attention on the intended message, beyond the learning contexts encountered.
Finally, all this learning cannot take place without safeguards. It is not a question of making students believe that they must blindly exercise a spirit of doubt and criticism on every piece of information. Such an attitude would quickly become completely counterproductive: by blindly doubting everything, whether from Internet users and bloggers or from our teacher or doctor, we are not adopting an effective critical mind. Teaching critical thinking therefore also includes education on how to deploy the acquired strategies wisely.

5.2 Do's

CT education - as defined in this Report - is based on three principles:

1. equip students' natural CT with knowledge and criteria appropriate to the context and progressively more sophisticated to assess the epistemic quality of the information at their disposal (plausibility, relevance, supporting evidence, quality of sources) in a more advanced, even expert way;
2. have students use their metacognitive confidence explicitly to calibrate the confidence that this information (or the decisions based on it) deserves, based on their assessment of the supporting evidence, the quality of the sources, and the relevance and plausibility of the information available;
3. provide students with as many elements as possible to arrive at this assessment and calibration. In addition to the criteria cited, help students:
    - gain a broader and deeper knowledge base to better assess the plausibility of information, but also to adjust confidence and limit metacognitive bias,

- identify of circumstances that are more likely to cause them difficulty and against which they must be particularly vigilant in their assessments and judgements. These circumstances motivate the use of criteria and explicit metacognitive thinking.

The considerations concerning the transfer of learning (presented in the fourth part) and others of a practical nature mean that the pedagogical approach proposed here leads students to exploit the content already taught and the skills worked on in different subjects, so that they understand, in an explicit way, that these disciplinary contents use strategies that can guide the evaluation of information. Indeed, curricula such as the one of the French National Education system largely cover the needs of teaching critical thinking, in terms of skills and content taught. What is missing, in our opinion, is:

- an assumed, explicit approach to the teaching of CT;
- a list of useful criteria for improving the evaluation of information ;
- practical indications for working on these criteria and knowledge in relation to CT, identifying the places in the curricula where these criteria and knowledge are already cited and present, without altering them (without introducing teaching hours dedicated to the teaching of critical thinking) but turning them in the direction of a CT education.

However, the indications presented here are not limited to the context of education in France: the principles can be adapted to any other education system, because of the contents and skills put forward by the latter.

The goal is to help students understand how certain criteria - introduced in science, history, literature, mathematics, sports, or other courses - are relevant to assessing the epistemic value of information (the evidence supporting that information or the source that carries the information), the quality of a source, or the relevance and plausibility of information, and how to mobilize them appropriately. Then invite them to place themselves on a continuum of trust in relation to the information being evaluated so that they can take a more reflective position.

5.2.1 Tooling up natural CT. How to choose the criteria to be taught?

A detailed list of criteria for expertly evaluating any type of information in any circumstance would probably be endless. In each area of knowledge, for each issue, experts in the field seek to develop ever finer criteria for evaluating information. Should we then give up teaching these criteria since it is impossible to make students expert on every issue? In reality, most of these expert criteria are unsuitable for everyday life situations and the content we come across. Primary and secondary education can therefore focus on criteria that meet two pragmatic conditions:

- they allow a more advanced and efficient evaluation of content regularly encountered in daily life or in the context of the knowledge provided by the school;
- they are easy to teach based on the disciplinary content and skills present in the school curriculum.

Thus, for example, every citizen, young and old, is exposed to assertions based on anecdotal, unsupported, chance observations. How much confidence should be placed in these assertions? This question carries considerable weight in everyday life. Every citizen is also exposed to and almost automatically fabricates causal explanations based on relatively general clues. How can it be established that a causal statement has a better chance of being correct than another? On the basis of which criteria and with the help of which tools can we get out of the uncertainty between two alternative hypotheses?

A starting point for answering these questions is what we do naturally, spontaneously, when we are confronted with new information, with a statement:

- we judge its plausibility, its consistency with the body of knowledge we have;
- We may assess the accompanying elements: these may be arguments that add to, supplement or justify the information and which we assess from a logical perspective in relation to their relevance to the information;
- we wonder about the reliability of the source who provides us with the information: its identity, its benevolence towards us, its lack of interest in lying or manipulating us (disinterested nature of the source); its competence in the field (expertise of the source);
- we may ask ourselves the question of how the evidence to support the claim was established: is it a direct observation? a memory? etc.

The objective of critical thinking education will be to build on these natural criteria of judgement (see Part 3 of this report) and to improve their use by making them more effective, more suitable for judging the truthfulness of an assertion or information. The aim is to achieve an advanced level of use of these criteria (spontaneous use being considered the basic level), and not an expert level, specific to certain professions (journalist, historian, professional scientist).

For this, it is a question of:

- impart knowledge to improve **plausibility** judgment ;
- learn how to identify a **good argument**, but especially the elements that can convince us wrongly (levers of persuasion, salient and available elements, desire to believe, desire to remove uncertainty...)
- learn how to identify **sources** that deserve our trust, in different contexts, based on the absence of private interests, and the competence of the source;
- learn about the recognition of better quality **evidence based on** the methods by which it was obtained (first-order epistemology).

5.2.2 Where can these criteria be found in curricula?

Where exactly are such criteria to be found in disciplinary teaching? Scientific teachings, for example, regularly propose to test a hypothesis (a misconception, an intuition based on observations, etc.). The teacher can present the task in the form of an evaluation of the merits of the idea. The teacher then sets up a procedure to judge whether the information in question is very reliable, plausible or probably false. The teacher may therefore choose to present the same lesson with an emphasis on the task that students have to do, i.e., to look for evidence in favour of the hypothesis (or its opposite). Scientific procedure is thus presented as a set of tools for producing evidence (and thus reliable knowledge).

Let's take a concrete example: we are in Vienna in the 19th century. In a hospital, an epidemic broke out among pregnant women, who were blamed in turn for the passage of the priest,

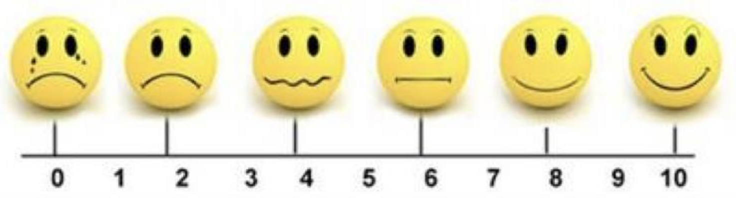
medical practices and the lack of windows. How to decide? The proposed pedagogical activity allows students to accumulate evidence in favour of each hypothesis and to acquire criteria to assess the reliability of the evidence collected. At the end, students must advise the hospital director of a decision based on the best available evidence.
Let us recognize that in daily life we often cannot mobilize such tools. However, we can use them as criteria to distinguish quality information from information that does not deserve our trust. Let's take another example: a scientific protocol is based on repeated observations because chance can produce particular events that we will then misinterpret. If we test the effect of a drug on a single patient, the patient may heal unrelated to the drug, but for many other reasons. Two or three patients are also not enough to produce reliable knowledge about the drug: it is necessary to multiply the observations. This tool of science can become a criterion for evaluating information for the citizen. Thus, a citizen confronted with contradictory assertions - one supported by a single personal testimony and the other supported by multiple observations - should trust the second assertion on the basis of the "multiple observations" criterion.

Such criteria are easy to find in science education in the narrower sense of the term (life and earth sciences, physical and chemical sciences) because the scientific method is an accumulation of tools to produce the best possible knowledge. The teacher can then stress the importance of multiplying observations, not confusing cause and correlation, varying only one factor at a time to look for a cause, etc. The teacher can then stress the importance of multiplying observations, not confusing cause and correlation, varying only one factor at a time to look for a cause, etc. Time for document analysis, investigation or debate is favourable for illustrating the concepts of facts, evidence and reliability. It is also important to stress the importance of making students understand the nature of the scientific enterprise for its own sake, as a social mechanism for the production of knowledge. The functioning of expert institutions that produce and evaluate knowledge is useful knowledge for any citizen to build informed confidence in scientific knowledge.

However, it is possible to address the importance of evaluating information on the basis of criteria in all disciplines:

- in mathematics, with the teaching of statistics and probability: the teacher can take advantage of all the situations where students study data sets and try to extract meaning from them to work on their perception of chance, the importance of multiplying observations, the use of more objective processing tools (such as averaging) while remaining vigilant about interpreting data in the absence of knowledge, etc. ;
- in technology, where the technological approach is closely associated with the scientific approach: the teacher can take every opportunity to use evidence-based hypotheses to point out the difference between mere opinions and reliable knowledge on which to base decisions;
- history and geography, other scientific disciplines where students regularly analyze the evidence and sources that support theses;
- in physical and sports education, where we may want to verify statements such as what we can do in sports and how to improve them;
- in literary teachings: it is possible to search in different texts for fact, description, proof, argument and interpretation. These texts have different purposes and it is essential to understand the author's intentions. The teacher can also compare the points of view of different cultures on the same phenomenon, or even on other cultures, to discuss the subjectivity of certain opinions, the origin of certain stereotypes, etc. ;
- in art education, students can question the role of the senses in perception. As in the study of texts, they can search in different works and media for the artist's intention to retranscribe reality (as in a naturalistic, anatomical drawing) or, on the contrary, to translate a subjective perception.

5.2.3 Exercising metacognitive trust explicitly. Place oneself on a continuum of trust and reflect on its meaning



Faced with information, we may feel more or less confident that its content is correct. This feeling can be explained, and we can reflect on what leads us to consider the information as reliable and trustworthy or, conversely, doubtful. The criteria we have presented above have this purpose. But we must make them operational, i.e. ensure that they influence our confidence in information, assertions and decisions. It is for this reason that we introduce here the concept, and the tool, of the trust continuum. Students reflect on the quality of evidence and the reliability of sources in support of a position. Even if the evidence is weak and the sources are not of exceptional quality, this does not prevent them from taking a position. However, they may reflect on the uncertainty that accompanies taking a position. As we discussed in the third part of this report, uncertainty is indeed a condition for a change of opinion, and recognizing the uncertainty that accompanies our positions can make it easier to understand and possibly adopt those of others. The continuum of trust therefore makes it easier to change our positions if necessary. It also makes it possible to express our certainty and confidence, and thus to show that certain opinions will not change so quickly. For example, standing on the right side of the continuum when faced with the statement "The Earth is rather spherical in shape" is a CT attitude, as long as the sources that assert it deserve our confidence and/or the arguments and evidence are within our reach.

5.2.4 Additional Elements for Achieving a Fair Assessment of Information Quality and Confidence Calibration

5.2.4.1 Acquiring a broader and deeper knowledge base

We have repeatedly stressed that CT is not divorced from the knowledge and content on which it operates. If, on the one hand, anchoring in content limits the transfer of the tools acquired to new contexts, on the other hand, a broader and deeper knowledge base would make it possible to better assess the plausibility of the information available. For example, many people would probably be prepared to react to the revelation that there is a dangerous liquid flowing through our faucets that can be fatal if inhaled directly, a liquid that is commonly used in the cooling systems of nuclear power plants, and is also the main component of acid rain: dihydrogen monoxide. The reactions would quickly be blurred by the knowledge that dihydrogen monoxide (H_2O) is actually water. A factual knowledge allows here to evaluate the scope of an information and in particular the plausibility of the assertions put forward concerning its risks. More generally, knowledge about energy and known forms of energy makes it possible to better assess the plausibility of claims about psychic energies or other claims that clearly contradict a large body of current knowledge in physics or biology. Possessing factual knowledge also increases the chances of being able to identify the most informed and reliable sources of information in a field, and this makes it possible to better calibrate our confidence and in particular to counteract the effect of metacognitive bias which manifests itself more particularly in relation to our zones of ignorance.

Instruction that teaches students a broad and connected picture of the main laws of nature and known physical and biological phenomena also provides them with a broad and solid basis for gauging the plausibility of a wealth of information. This does not mean that other information for which the knowledge developed by the individual is weak will not be judged more naively. The dependence of CT on knowledge content inevitably makes us fallible. The aim is to improve the capabilities of CT by equipping it, not to achieve ideal perfection.

5.2.4.2 Become aware of situations where advanced CT is required and of circumstances that are more likely to put us at risk

The deployment of advanced criteria for assessing information quality and the voluntary and explicit exercise of metacognitive trust judgements requires effort and comes at a cost compared to the use of more intuitive criteria that do not involve the acquisition of knowledge or implicit assessments.

We should not give the impression that the mobilization of these criteria is a constant necessity, and that every moment of life is going to turn into a heavy exercise in reflection. These conditions are mobilized only when the situation requires it: a new content, an important decision to be taken, an opinion which, once fixed, may have not insignificant consequences for oneself or for others.

However, motivating the use of such criteria is also a challenge. Indeed, we know that multiple observations are more reliable than limited observations, or that hidden sources are suspect. But we also have at our disposal intuitive evaluation strategies, often less costly, especially in busy "ecological" contexts such as those we encounter in our daily lives. Sometimes we have to make decisions quickly, in situations where other issues are involved (the gaze of others, emotions, etc.). We should therefore not underestimate the difficulty of the effort we must make to make the use of these criteria accessible. A lever to play on therefore concerns the explanation of the limits of our natural tools. The teacher can create pedagogical situations that lead students to respond intuitively to a problem in order to detect the flaws of such strategies. For example, if several groups of students reach different conclusions after a limited sampling, the teacher can rebound by showing them the importance of not giving in to the natural temptation to conclude from limited observations. This reinforces the alternative, more costly strategy (relying on multiple observations).

The value of relying on concrete situations is clear here: rather than conveying abstract discourse and general advice that might seem trivial ("repeated observations are better than isolated observation"), the pedagogical situation encourages real reflection by the student on his or her own strategies and on the methods that produce reliable or unreliable conclusions. These situations rooted in realistic contexts for the child are more likely to enable him or her to develop information assessment skills if they are subsequently associated with more abstract discourse ("conclusions from repeated observations are more reliable in principle than those based on one or two observations"). This abstract formulation is more likely to be memorized and could be mobilized later in other contexts. However, it would only make sense if it had been previously associated with learning in a context that is relevant to the student.

5.2.5 Testing

In this section, we have sought to provide general principles for CT education that we believe are derived from the literature reviews and analyses presented in this report. No single principle, although based on documented knowledge, can guarantee the effectiveness of an instructional approach under ecological, classroom conditions. The knowledge we have mobilized nevertheless seems to us to justify the effort and opportunity to test the effectiveness of educational interventions that draw on it. Empirical data from the field can tell us more about how to approach CT education in the most effective way possible.

Annexes

Notes

---

[1] *"Critical thinking is a widely accepted educational goal. Its definition is contested, but the competing definitions can be understood as differing conceptions of the same basic concept: careful thinking directed to a goal. Conceptions differ with respect to the scope of such thinking, the type of goal, the criteria and norms for thinking carefully, and the thinking components on which they focus. Its adoption as an educational goal has been recommended on the basis of respect for students' autonomy and preparing students for success in life and for democratic citizenship. 'Critical thinkers' have the dispositions and abilities that lead them to think critically when appropriate. The abilities can be identified directly; the dispositions indirectly, by considering what factors contribute to or impede exercise of the abilities. Standardized tests have been developed to assess the degree to which a person possesses such dispositions and abilities. Educational intervention has been shown experimentally to improve them, particularly when it includes dialogue, anchored instruction, and mentoring. Controversies have arisen over the generalizability of critical thinking across domains, over alleged bias in critical thinking theories and instruction, and over the relationship of critical thinking to other types of thinking."*

Hitchcock 2018

[2] "*It is daunting for someone who is becoming acquainted with the field of critical thinking to find the assortment of definitions provided by long-time participants in the field. It might seem from this assortment of definitions that the field of critical thinking is in chaos. But I do not think it is. What seems more reasonable to me is that the set of definitions in Table 1 below are different descriptions of the same concept… which I hold is the mainstream concept of critical thinking. ...*

These are definitions with which over the years I have become familiar while pursuing my interests in critical thinking."
*1) "**Active, persistent, and careful consideration** of any belief or supposed form of knowledge in the light of the **grounds that support it** and the further conclusions to which it tends" (Dewey, 1933, p. 9 (first edition 1910).*
*2) "The ability to think critically ...involves three things: (1) an attitude of being disposed to **consider in a thoughtful way** the problems and subjects that come within the range of one's experiences, (2) knowledge of the methods of logical inquiry and reasoning, and (3) some skill in applying those methods (Glaser 1941).*
*3) "Critical thinking is the intellectually disciplined process of **actively and skillfully conceptualizing, applying, analyzing, synthesizing,** and/or **evaluating information** gathered from, or generated by, observation, experience, reflection, reasoning, or communication, as a guide to belief and action" (Scriven and Paul 1987).*
*4) "Critical thinking is reasonable **reflective thinking** that is focused on **deciding what to believe or** do" (Ennis 1987a; 1987b; 1991; 2011, 2015).*
*5) "A critical thinker is one who is appropriately moved by **reasons**" (Siegel 1988, p. 32).*
*6) "Skillful, responsible thinking that is conducive to good judgment because it is sensitive to context, **relies on criteria, and is self-correcting**" (Lipman 1988).*
*7) "**Purposeful**, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual **considerations upon which that judgment is based"** (Facione 1990, Table 1).*
*8) "Critical thinking is skilled, **active** interpretation and **evaluation of observations, communications, information, and argumentatio**n as a guide to thought and action" (Fisher and Scriven 1997, p. 20)1.*
*9) "The practice of **identifying, having, and giving good reasons** for one's beliefs, values, and actions, given one's goals of truth and avoidance of error" (Possin, 2002).*
*10) "Thinking that attempts to arrive at a judgment only after **honestly evaluating** alternatives with respect to available evidence and arguments" (Hatcher and Spencer, 2006, p. 1).*
*11) "The **careful examination** of an issue in order to reach a reasoned judgment" (Bailin and Battersby, 2010).*

---

12) "The ***careful, deliberate determination of whether we should accept, reject, or suspend judgment*** *about a claim, and the degree of confidence with which we accept or reject it" (Moore and Parker, 2014).*
13) "The ***articulated judgment*** *of an intellectual product arrived at on the basis of plus-minus considerations of the product in terms of appropriate standards (or criteria)" (Johnson, 2014).*
*14) "[****Reasonable****] inquiry and argument" (Kuhn, 2015, p. 47)*

"Non-academic internet-dictionary definitions of 'critical thinking', which I assume have captured current use of the term in the media, also express the concept of critical thinking."

1) "Disciplined thinking that is clear, rational, open-minded, and informed by evidence" (Dictionary.com, 2016)
2) "The objective analysis and evaluation of an issue in order to form a judgment"(Oxforddictionaries.com, 2016)
3) "A mental process of reviewing clear, rational thoughts based on evidence to reach an answer or a conclusion (Yourdictionary.com, 2016)."

Ennis 2016

[3] "*Although many psychologists and others have proposed definitions for the term 'critical thinking', these definitions tend to be similar with considerable overlap among definitions. In a review of the critical thinking literature, Fischer & Spiker (2000) found that most definitions for the term 'critical thinking' include reasoning/logic, judgment, metacognition, reflection, questioning and mental processes. Jones and his colleagues (Jones, et al. 1995a, b) obtained consensus from among 500 policy makers, employers, and educators who agree that critical thinking is a broad term that describes reasoning in an open-ended manner and with an unlimited number of solutions. It involves constructing a situation and supporting the reasoning that went into a conclusion."*

Halpern 2013

[4] "*Judging by the attention given to critical thinking in educational journals and in the social documents of governing agencies, support for teaching critical thinking at all levels of education is extremely strong in North America and the UK. But, agreement about teaching critical thinking persists only so long as theorists remain at the level of abstract discussion and permit their use of the term to remain vague. As soon as they begin to spell out in more concrete terms what critical thinking consists in, what educational attainments are required if one is to be a critical thinker, and what means are likely to be efficacious in teaching persons to think critically, that is to say, as soon as they interpret the term in such a way as to provide a clear conception of critical thinking, agreement evaporates*."

Bailin et al. 1999

[5] "*The experts articulated an ideal. It may be that no person is fully adept at the skills and sub-skills the experts found to be central to CT. It may be that no person has fully cultivated all the affective dispositions which characterize a good critical thinker. It may be that humans compartmentalize their lives in ways that CT is more active and evident in some areas than in others. This gives no more reason to abandon the effort to infuse CT into the educational system than that knowing no friendship is perfect gives reason to despair of having friends. The experts' proposal in putting the ideal before education community is that it should serve as a rich and worthy goal guiding CT assessment and curriculum development at all educational levels. We understand critical thinking to be the purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based. CT is essential as a tool of inquiry. As such, CT is a liberating force in education and a powerful resource in one's personal and civic life. While not synonymous with good thinking, CT is a pervasive and self-rectifying human phenomenon. The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit. Thus, educating good critical thinkers means working toward this ideal. It combines developing CT skills with nurturing those dispositions which consistently yield useful insights and which are the basis of a rational and democratic society."*

Facione 1990

[6] "*Individuals vary in their views of what students should be taught … But there is no disagreement on the importance of critical thinking skills. In free societies, the ability to think critically is viewed as a cornerstone of individual civic engagement and economic success. We may disagree about which content students should learn, but we at least agree that, whatever they end up learning, students ought to think critically about it. … Despite this consensus it is not clear we know what we mean by 'critical thinking'. I will offer a commonsensical view (Willingham, 2007). You are thinking critically if (1) your thinking is novel - that is, you aren't simply drawing a conclusion from a memory of a previous situation and (2) your thinking is self-directed - that is, you are not merely executing instructions given by someone else and (3) your thinking is effective - that is, you respect certain conventions that make thinking more likely to yield useful conclusions. These would be conventions like 'consider both sides of an issue', and 'offer evidence for claims made', and 'don't let emotion interfere with reason'. This last characteristic will be our main concern, and as we'll see, what constitutes effective thinking varies from domain to domain.*"

Willingham 2019

[7] "*The intellectual roots of critical thinking are as ancient as its etymology, traceable, ultimately, to the teaching practice and vision of Socrates 2,500 years ago who discovered by a method of probing questioning that people could not rationally justify their confident claims to knowledge. Confused meanings, inadequate evidence, or self-contradictory beliefs often lurked beneath smooth but largely empty rhetoric. Socrates established the fact that one cannot depend upon those in 'authority' to have sound knowledge and insight. He demonstrated that persons may have power and high position and yet be deeply confused and irrational. He established the importance of asking deep questions that probe profoundly into thinking before we accept ideas as worthy of belief. He established the importance of seeking evidence, closely examining reasoning and assumptions, analyzing basic concepts, and tracing out implications not only of what is said but of what is done as well. His method of questioning is now known as 'Socratic Questioning' and is the best known critical thinking teaching strategy. In his mode of questioning, Socrates highlighted the need in thinking for clarity and logical consistency. …*

Socrates' practice was followed by the critical thinking of Plato (who recorded Socrates' thought), Aristotle, and the Greek skeptics, all of whom emphasized that things are often very different from what they appear to be and that only the trained mind is prepared to see through the way things look to us on the surface (delusive appearances) to the way they really are beneath the surface (the deeper realities of life). From this ancient Greek tradition emerged the need, for anyone who aspired to understand the deeper realities, to think systematically, to trace implications broadly and deeply, for only thinking that is comprehensive, well-reasoned, and responsive to objections can take us beyond the surface. …

*In the Middle Ages, the tradition of systematic critical thinking was embodied in the writings and teachings of such thinkers as Thomas Aquinas (**Sumna Theologica**) who to ensure his thinking met the test of critical thought, always systematically stated, considered, and answered all criticisms of his ideas as a necessary stage in developing them. … In the Renaissance (15th and 16th Centuries), a flood of scholars in Europe began to think critically about religion, art, society, human nature, law, and freedom. … Francis Bacon, in England, was explicitly concerned with the way we misuse our minds in seeking knowledge. He recognized explicitly that the mind cannot safely be left to its natural tendencies. In his book **The Advancement of Learning**, he argued for the importance of studying the world empirically. He laid the foundation for modern science with his emphasis on the information-gathering processes. He also called attention to the fact that most people, if left to their own devices, develop bad habits of thought (which he called 'idols') that lead them to believe what is false or misleading. ...*

*Some fifty years later in France, Descartes wrote what might be called the second text in critical thinking, **Rules For the Direction of the Mind**. In it, Descartes argued for the need for a special systematic disciplining of the mind to guide it in thinking. He articulated and defended the need in thinking for clarity and precision. He developed a method of critical thought based on the principle of systematic doubt. …*

Hobbes and Locke (in 16th and 17th Century England) displayed the same confidence in the critical mind of the thinker that we find in Machiavelli. Neither accepted the traditional picture of things dominant in the thinking of their day. Neither accepted as necessarily rational that which was considered 'normal' in their culture. Both looked to the critical mind to open up new vistas of learning. Hobbes adopted a naturalistic view of the world in *which everything was to be explained by evidence and reasoning. Locke defended a common-sense analysis of everyday life and thought. He laid the theoretical foundation for critical thinking about basic human rights and the responsibilities of all governments to submit to the reasoned criticism of thoughtful citizens. It was in this spirit of intellectual freedom and critical thought that people such as Robert Boyle (in the 17th Century) and Sir Isaac Newton (in the 17th and 18th Century) did their work. In his* ***Sceptical Chymist****, Boyle severely criticized the chemical theory that had preceded him. Newton, in turn, developed a far-reaching framework of thought which roundly criticized the traditionally accepted world view. He extended the critical thought of such minds as Copernicus, Galileo, and Kepler. After Boyle and Newton, it was recognized by those who reflected seriously on the natural world that egocentric views of world must be abandoned in favor of views based entirely on carefully gathered evidence and sound reasoning. …*

Another significant contribution to critical thinking was made by the thinkers of the French Enlightenment: Bayle, Montesquieu, Voltaire, and Diderot. They all began with the premise that the human mind, when disciplined by reason, is better able to figure out the nature of the social and political world. What is more, for these thinkers, reason must turn inward upon itself, in order to determine weaknesses and strengths of thought. They valued disciplined intellectual exchange, in which all views had to be submitted to serious analysis and critique. They believed that all authority must submit in one way or another to the scrutiny of reasonable critical questioning."

Foundation for critical thinking, https://www.criticalthinking.org/pages/defining-critical-thinking/766

[8] "*The essence of critical thinking is suspended judgment; and the essence of this suspense is inquiry to determine the nature of the problem before proceeding to attempts at its solution. This, more than any other thing, transforms mere inference into tested inference, suggested conclusions into proof. ...*

Active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends, constitutes reflective thought."

Dewey 1933

[9] "*The real problem of intellectual education is the transformation of more or less casual curiosity and sporadic suggestion into attitudes of alert, cautious, and thorough inquiry."*

Dewey 1933

"Thinking is stoppage of the immediate manifestation of impulse until that impulse has been brought into connection with other possible tendencies to action so that a more comprehensive and coherent plan of activity is formed. Some of the other tendencies to action lead to use of eye, ear, and hand to observe objective conditions; others result in recall of what has happened in the past. Thinking is thus a postponement of immediate action, while it effects internal control of impulse through a union of observation and memory, this union being the heart of reflection. What has been said explains the meaning of the well-worn phrase 'self-control'. The ideal aim of education is creation of power of self-control."

Dewey 1997

[10] *1. Transit: "The other day, when I was downtown on 16th Street, a clock caught my eye. I saw that the hands pointed to 12:20. This suggested that I had an engagement at 124th Street, at one o'clock. I reasoned that as it had taken me an hour to come down on a surface car, I should probably be twenty minutes late if I returned the same way. I might save twenty minutes by a subway express. But was there a station near? If not, I might lose more than twenty minutes in looking for one. Then I thought of the elevated, and I saw there was such a line within two blocks. But where was the station? If it were several blocks above or below the street I was on, I should lose time instead of gaining it. My mind went back to the subway express as quicker than the elevated; furthermore, I remembered that it went nearer than the elevated to the part of 124th Street I wished to reach, so that time would be saved at the end of the journey. I concluded in favor of the subway, and reached my destination by one o'clock."*

2. Ferryboat: "Projecting nearly horizontally from the upper deck of the ferryboat on which I daily cross the river is a long white pole, bearing a gilded ball at its tip. It suggested a flagpole when I first saw it; its color, shape, and gilded ball agreed with this idea, and these reasons seemed to justify me in this belief. But soon *difficulties presented themselves. The pole was nearly horizontal, an unusual position for a flagpole; in the next place, there was no pulley, ring, or cord by which to attach a flag; finally, there were elsewhere two vertical staffs from which flags were occasionally flown. It seemed probable that the pole was not there for flag-flying. I then tried to imagine all possible purposes of such a pole, and to consider for which of these it was best suited: (a) Possibly it was an ornament. But as all the ferryboats and even the tugboats carried poles, this hypothesis was rejected. (6) Possibly it was the terminal of a wireless telegraph. But the same considerations made this improbable. Besides, the more natural place for such a terminal would be the highest part of the boat, on top of the pilot house, (c) Its purpose might be to point out the direction in which the boat is moving. In support of this conclusion, I discovered that the pole was lower than the pilot house, so that the steers man could easily see it. Moreover, the tip was enough higher than the base, so that, from the pilot's position, it must appear to project far out in front of the boat. Moreover, the pilot being near the front of the boat, he would need some such guide as to its direction. Tug boats would also need poles for such a purpose. This hypothesis was so much more probable than the others that I accepted it. I formed the conclusion that the pole was set up for the purpose of showing the pilot the direction in which the boat pointed, to enable him to steer correctly."*

3. Bubbles: "In washing tumblers in hot soap suds and placing them mouth downward on a plate, I noticed that bubbles appeared on the outside of the mouth of the tumblers and then went inside. Why? The presence of bubbles suggests air, which I note must come from inside the tumbler. I see that the soapy water on the plate prevents escape of the air save as it may be caught in bubbles. But why should air leave the tumbler? There was no substance entering to force it out. It must have expanded. It expands by increase of heat or by increase of pressure, or by both. Could the air have become heated after the tumbler was taken from the hot suds? Clearly not the air that was already entangled in the water. If heated air was the cause, cold air must have entered in transferring the tumblers from the suds to the plate. I test to see whether this supposition is true by taking several more tumblers out. Some I shake so as to make sure of entrapping cold air in them. Some I take out, holding them mouth downward in order to prevent cold air from entering. Bubbles appear on the outside of every one of the former and on none of the latter. I must be right in my inference. Air from the outside must have been expanded by the heat of the tumbler, which explains the appearance of the bubbles on the outside. But why do they then go inside? Cold contracts. The tumbler cooled and also the air inside it. Tension was removed, and hence bubbles appeared inside. To be sure of this, I test by placing a cap of ice on the tumbler while the bubbles are still forming outside. They soon reverse."

Dewey 1933

[11] "*1. An attitude of being disposed to consider in a thoughtful way the problems and subjects to come within the range of one's own experiences; 2. Knowledge of the methods of logical inquiry and reasoning; 3. Some skills in applying those methods. Critical thinking calls for a persistent effort to examine any belief or supposed form of knowledge in the light of the evidence that supports it and the further conclusions to which it tends. It also generally requires ability to recognise problems, to find workable means for meeting this problems, to gather and marshal pertinent information, to recognise unstated assumptions and values, to comprehend and use language with accuracy, clarity and discrimination, to interpret data, to appraise evidence and evaluate arguments, to recognise the existence (or non-existence) of logical relationships between propositions, to draw warranted conclusions and generalizations at which one arrives, to reconstruct one's patterns of belief on the basis of wider experience, and to render accurate judgments about specific things and qualities in everyday life."*

Glaser 1941

[12] "*The application of critical thought to pedagogy in our schools was given a major impetus in the middle of this century. According to Cuban (1984), 'The work of B.O. Smith in the 1950s and subsequently Robert Ennis, have provided a scholarly rationale and specific ingredients for designing school programs to develop critical thought'."*

Lewis & Smith 1993

[13] "*The Philosophy for Children Program, developed by Matthew Lipman, represents one way to introduce critical thinking skills. … In Lipman's program, fifth and sixth grade children read and discuss a set of novels. Through these activities, students are encouraged to develop philosophical reasoning skills including commitments to impartiality and objectivity, relevance, consistency, and the search for defensible reasons for behavior. While Lipman's program is designed especially to be added to the curriculum, other philosophers*

incorporate critical thinking into the existing curriculum. For example, Paul, Binker, & Weil's Critical Thinking Handbook (1990) helps K-3 teachers remodel their lesson plans in language arts, social studies, and science in order to incorporate critical thinking."

Lewis & Smith 1993

[14] "*Critical thinking is a cultivation of that strand of traditional education which stresses the cultivation of wisdom and its application to both practice and life. To strengthen critical thinking in schools and colleges, it is necessary to know its defining features, its characteristic outcomes, and its underlying conditions. The outcomes of critical thinking are judgments; and the nature of judgment is such that critical thinking may be defined as skillful, responsible thinking that facilitates good judgment because (1) it relies upon criteria; (2) it is self-correcting, and (3) it is sensitive to context. The very meaning of 'criterion' is 'a rule or principle utilized in the making of judgments.' Judgment, in turn, is a skill; therefore critical thinking is skillful thinking, and skills can only be defined through criteria by which performance can be evaluated. So critical thinking is thinking that both employs criteria and can be assessed by appeal to criteria. Important criteria are reliability, strength, relevance, coherence, and consistency. Critical thinking is self-corrective, promoting a community of inquiry in the classroom by requiring students to discover weaknesses in their own thinking and to rectify faults in their procedures. Finally, thinking that is sensitive to context involves recognition of: exceptional or irregular circumstances and conditions; special limitations, contingencies, or constraints; overall configurations; the possibility that evidence is atypical; and the possibility that some meanings do not translate from one context or domain to another. Exemplary instances of critical thinking can be found in the best practice of law and medicine. The relevance of critical thinking to the enhancement of K-12 and college education is related to the shift from learning to thinking as the focus of education and to the goal of helping students develop the reasoning skills that will enable them to exercise good judgment.'*

Lipman, 1988

[15] "*It is true of teachers and farmers and theoretical physicists as well: all must make judgments in the practice of their occupations and in the conduct of their lives. There are practical, productive, and theoretical judgments, as Aristotle would have put it. Insofar as we make such judgments well, we can be said to behave wisely.(...) They are likely to be good judgments if they are the product of skillfully performed acts guided or facilitated by appropriate instruments and procedures."*

Lipman 1987

"Thus architects will judge a building by using criteria such as utility, safety, and beauty. And presumably critical thinkers will rely upon such time-tested criteria such as validity, evidential warrant, and consistency. Any area of practice - architectural, cognitive and the like - should be able to cite the criteria by which the practice is guided."

Lipman 1987

[16] *"The eighties witnessed a growing accord that the heart of education lies exactly where traditional advocates of a liberal education always said it was — in the processes of inquiry, learning and thinking rather than in the accumulation of disjointed skills and senescent information. By the decade's end the movement to infuse the K-12 and post-secondary curricula with critical thinking (CT) had gained remarkable momentum. This success also raised vexing questions: What exactly are those skills and dispositions which characterize CT? What are some effective ways to teach CT? And how can CT, particularly if it becomes a campus-wide, district-wide or statewide requirement, be assessed? When asked by the individual professor or teacher seeking to introduce CT into her own classroom, such questions are difficult enough. But they take on social, fiscal, and political dimensions when asked by campus curriculum committees, school district offices, boards of education, and the educational testing and publishing industries. Given the central role played by philosophers in articulating the value, both individual and social, of CT, in analyzing the concept of CT, in designing college level academic programs in CT, and in assisting with efforts to introduce CT into the K-12 curriculum, it is little wonder that the American Philosophical Association, through its Committee on Pre-College Philosophy, took great interest in the CT movement and its impact on the profession. In December of 1987 that committee asked this investigator to make a systematic inquiry into the current state of CT and CT assessment."*

Facione 1990

[17] "*At one level we all know what 'critical thinking' means — it means good thinking, almost the opposite of illogical, irrational, thinking. But when we test our understanding further, we run into questions. For example, is critical thinking the same as creative thinking, are they different, or is one part of the other? How do critical thinking and native intelligence or scholastic aptitude relate? Does critical thinking focus on the subject matter or content that you know or on the process you use when you reason about that content? ...*

Suggestion: What can the strong critical thinkers do (what mental abilities do they have), that the weak critical thinkers have trouble doing? What skills or approaches do the strong critical thinkers habitually seem to exhibit which the weak critical thinkers seem not to possess?

...look for a list of mental skills and habits of mind, the experts, when faced with the same problem you are working on, refer to their lists as including cognitive skills and dispositions. As to the cognitive skills here is what the experts include as being at the very core of critical thinking: interpretation, analysis, evaluation, inference, explanation, and self-regulation. …

Quoting from the consensus statement of the national panel of experts: ***interpretation*** *is 'to comprehend and express the meaning or significance of a wide variety of experiences, situations, data, events, judgments, conventions, beliefs, rules, procedures, or criteria.' Interpretation includes the sub-skills of categorization, decoding significance, and clarifying meaning. … analysis is 'to identify the intended and actual inferential relationships among statements, questions, concepts, descriptions, or other forms of representation intended to express belief, judgment, experiences, reasons, information, or opinions.' The experts include examining ideas, detecting arguments, and analyzing arguments as sub-skills of analysis. …*

evaluation *as meaning 'to assess the credibility of statements or other representations which are accounts or descriptions of a person's perception, experience, situation, judgment, belief, or opinion; and to assess the logical strength of the actual or intended inferential relationships among statements, descriptions, questions or other forms of representation.' … To the experts* ***inference*** *means 'to identify and secure elements needed to draw reasonable conclusions; to form conjectures and hypotheses; to consider relevant information and to deduce the consequences flowing from data, statements, principles, evidence, judgments, beliefs, opinions, concepts, descriptions, questions, or other forms of representation.' As sub-skills of inference the experts list querying evidence, conjecturing alternatives, and drawing conclusions.*

Beyond being able to interpret, analyze, evaluate and infer, strong critical thinkers can do two more things. They can explain what they think and how they arrived at that judgment. And, they can apply their powers of critical thinking to themselves and improve on their previous opinions. These wo skills are called 'explanation' and 'self-regulation.' The experts define ***explanation*** *as being able to present in a cogent and coherent way the results of one's reasoning. This means to be able to give someone a full look at the big picture: both 'to state and to justify that reasoning in terms of the evidential, conceptual, methodological, criteriological, and contextual considerations upon which one's results were based; and to present one's reasoning in the form of cogent arguments.' The sub-skills under explanation are describing methods and results, justifying procedures, proposing and defending with good reasons one's causal and conceptual explanations of events or points of view, and presenting full and well- reasoned, arguments in the context of seeking the best understandings possible. …*

Maybe the most remarkable cognitive skill of all, however, is this next one. This one is remarkable because it allows strong critical thinkers to improve their own thinking. In a sense this is critical thinking applied to itself. Because of that some people want to call this 'meta-cognition,' meaning it raises thinking to another level. But 'another level' really does not fully capture it, because at that next level up what self-regulation does is look back at all the dimensions of critical thinking and double check itself. … 'self-consciously to monitor one's cognitive activities, the elements used in those activities, and the results educed, particularly by applying skills in analysis, and evaluation to one's own inferential judgments with a view toward questioning, confirming, validating, or correcting either one's reasoning or one's results.' The two sub- skills here are self-examination and self-correction."

Facione 2011

[18] *"What kind of a person would be apt to use their critical thinking skills? The experts poetically describe such a person as having 'a critical spirit.' Having a critical spirit does not mean that the person is always negative and hypercritical of everyone and everything. The experts use the metaphorical phrase critical spirit in a positive sense. By it they mean 'a probing inquisitiveness, a keenness of mind, a zealous dedication to reason,*

and a hunger or eagerness for reliable information. The kind of person being described here is the kind that always wants to ask 'Why?' or 'How?' or 'What happens if?'. The one key difference, however, is that in fiction Sherlock always solves the mystery, while in the real world there is no guarantee. Critical thinking is about how you approach problems, questions, issues. It is the best way we know of to get to the truth. … The approaches to life and living which characterize critical thinking include:

** inquisitiveness with regard to a wide range of issues,
* concern to become and remain well-informed,
* alertness to opportunities to use critical thinking,
* trust in the processes of reasoned inquiry,
* self-confidence in one's own abilities to reason,
* open-mindedness regarding divergent world-views,
* flexibility in considering alternatives and opinions,
* understanding of the opinions of other people,
* fair-mindedness in appraising reasoning,
* honesty in facing one's own biases, prejudices, stereotypes, or egocentric tendencies,
* prudence in suspending, making or altering judgments,
* willingness to reconsider and revise views where honest reflection suggests that change is warranted.*

What would someone be like who lacked those dispositions? It might be someone who does not care about much of anything, is not interested in the facts, prefers not to think, mistrusts reasoning as a way of finding things out or solving problems, holds his or her own reasoning abilities in low esteem, is close-minded, inflexible, insensitive, cannot understand what others think, is unfair when it comes to judging the quality of arguments, denies his or her own biases, jumps to conclusions or delays too long in making judgments, and never is willing to reconsider an opinion. … The experts went beyond approaches to life and living in general to emphasize that strong critical thinkers can also be described in terms of how they approach specific issues, questions, or problems. The experts said you would find these sorts of characteristics: approaches to life and living in general to emphasize that strong critical thinkers can also be described in terms of how they approach specific issues, questions, or problems. The experts said you would find these sorts of characteristics:

** clarity in stating the question or concern,
* orderliness in working with complexity,
* diligence in seeking relevant information,
* reasonableness in selecting and applying criteria,
* care in focusing attention on the concern at hand,
* persistence though difficulties are encountered,
* precision to the degree permitted by the subject and the circumstances."*

Facione 2011

[19] "*One system is more intuitive, reactive, quick and holistic. So as not to confuse things with the notions of thinking in popular culture, cognitive scientists often name this system, 'System 1.' The other (yes, you can guess its name) is more deliberative, reflective, computational and rule governed. You are right, it is called 'System 2'. In* ***System 1*** *thinking, one relies heavily on a number of heuristics (cognitive maneuvers), key situational characteristics, readily associated ideas, and vivid memories to arrive quickly and confidently at a judgment. System 1 thinking is particularly helpful in familiar situations when time is short and immediate action is required. While System 1 is functioning, another powerful system is also at work, that is, unless we shut it down by abusing alcohol or drugs, or with fear or indifference. Called '***System 2***,' this is our more reflective thinking system. It is useful for making judgments when you find yourself in unfamiliar situations and have more time to figure things out. It allows us to process abstract concepts, to deliberate, to plan ahead, to consider options carefully, to review and revise our work in the light of relevant guidelines or standards or rules of procedure. While System 2 decisions are also influenced by the correct or incorrect application of heuristic maneuvers, this is the system which relies on well-articulated reasons and more fully developed evidence. It is reasoning based on what we have learned through careful analysis, evaluation, explanation, and self-correction. This is the system which values intellectual honesty, analytically anticipating what happens next, maturity of judgment, fair-mindedness, elimination of biases, and truth-seeking. This is the system which we rely on to think carefully through complex, novel, high-stakes, and highly integrative problems … Educators urge us to improve our critical thinking skills and to reinforce our disposition to use those skills because that is perhaps the best way to*

develop and refine our System 2 reasoning. System 1 and System 2 are both believed to be vital decision-making tools when stakes are high and when uncertainty is an issue. Each of these two cognitive systems are believed to be capable of functioning to monitor and potentially override the other. This is one of the ways our species reduces the chance of making foolish, sub-optimal or even dangerous errors in judgment. Human thinking is far from perfect. Even a good thinker makes both System 1 and 2 errors. At times we misinterpret things, or we get our facts wrong, and we make mistakes as a result. But often our errors are directly related to the influences and misapplications of cognitive heuristics."

Facione 2011

[20] "*Any defensible conception must construe critical thinking in such a way as to capture most of what people have in mind when they claim that developing critical thinking is an important goal of education. That is to say, it must be true to the core meaning of the educator's basic concept of critical thinking. Should it fail in this regard, it is largely irrelevant to educators concerned with developing critical thinking. What, then, do critical thinking advocates generally have in mind when they talk about critical thinking? ... thinking regarded as critical thinking must be directed toward some end or purpose, such as answering a question, making a decision, solving a problem, resolving an issue, devising a plan, or carrying out a project. Roughly speaking, thinking that serves such purposes can be characterized as thinking aimed at forming a judgement, i.e. making up one's mind about what to believe or do. ... Not just any thinking aimed at deciding what to believe or do can count as critical thinking, however. If the thinking is sloppy, superficial, careless, rash or naive, most advocates of critical thinking would not agree that it is critical thinking. For example, someone who comes to believe on the basis of poor or irrelevant reasons, on the authority of someone whose credibility is questionable, or without attempting to assess the evidence relevant to the truth of the belief, would not usually be regarded as thinking critically. This suggests that thinking about what to believe or do must meet appropriate standards if it is to be regarded as critical thinking. Moreover, these standards cannot be met merely by accident or happenstance. If someone were inadvertently to fulfill relevant standards in their thinking, but had not intentionally attempted to fulfill them, they would not generally be regarded as having engaged in critical thinking. To be engaged in critical thinking one must be aware that there are such standards and must be striving to fulfill them. This is not to say of course, that a person engaged in thinking critically is necessarily able to state or verbally explicate the relevant standards. To summarize, critical thinking, as it is typically understood by educators, has at least these three features:*

* it is done for the purpose of making up one's mind about what to believe or do,
* *the person engaging in the thinking is trying to fulfill standards of*
* *the adequacy and accuracy appropriate to the thinking and the thinking fulfill the relevant standards to some threshold level."*

Bailin et al. 1999

[21] "*...this construal is too narrow, in that it fails to do justice to the fact that critical thinking very often takes place in the context of persons' thinking things through together by means of discussion and dialogue. Popper (1972: 148) has emphasized the importance of critical discussion in the advancement of science. We believe it plays an equally important role in most areas of inquiry and practice, including political and moral decision-making. If we are correct in supposing that group deliberation is an important context for critical thinking, then the thinking appropriate to such contexts must be included in our conception of critical thinking. This means that, in addition to assessing intellectual products appropriately, critical thinking will include responding constructively to reasons and arguments given by others in the context of discussion. Responding constructively in such a context involves furthering the point or purpose of the critical discussion, while maintaining a social environment that enables all parties to the discussion to participate fully. Thus, good thinking in this context involves more than good reason-evaluation."*

Bailin et al. 1999

[22] "*Because critical thinking is, in our view, thinking in such a way as to fulfill relevant standards, it is the standards of good thinking that provide the criteria for determining what attributes are important for critical thinkers. If an attribute is required by persons in order to fulfill a standard of good thinking, or if it will significantly increase the chances that their thinking fulfill such standards, it can legitimately be regarded as an attribute that should be fostered in a critical thinker."*

Bailin et al. 1999

[23] "*It is fairly common to characterize the critical thinker by enumerating a list of skills or abilities and a list of attitudes or dispositions such a thinker must have. This kind of characterization is appealing, because obviously there are certain kinds of things critical thinkers must be able and inclined to do. However, adoption of such language, and in particular the use of the terms 'skills' and 'abilities', has the potential for causing confusion. The lists of critical thinking skills or abilities offered by critical thinking theorists are necessarily a list of things the critical thinker must be able to accomplish, for the only way we have of describing what one is able to do in thinking is in terms of the outcomes generated by the thinking. Thus, we say that the critical thinker must be able to do such things as judge the adequacy of reportive definitions, detect invalid arguments, etc. … Such lists imply nothing about the psychological states, capacities or processes that enable critical thinkers to have the requisite accomplishments, and nothing about the kinds of instructional procedures that are likely to be efficacious in bringing them about. Misunderstanding arises, however, when we begin to describe the items on such a list as 'critical thinking abilities or skills', that is, when we begin to talk about the ability to judge reportive definitions' or 'the skill of detecting invalid arguments.' Many educators interpret such ability and skill descriptions as descriptions of psychological processes, states or capacities, rather than as simply descriptions of what persons can accomplish."*

Bailin et al. 1999

[24] "*To a considerable extent, the quality of thinking persons are able to do about a particular problem, issue or question is determined by what they know, or are able to find out, about it and about the context in which it must be resolved. Moreover, critical thinking always takes place in the context of (and against the backdrop of) already existing concepts, beliefs, values, and ways of acting. This context plays a very significant role in determining what will count as sensible or reasonable application of standards and principles of good thinking. Thus, the depth of knowledge, understanding and experience persons have in a particular area of study or practice is a significant determinant of the degree to which they are capable of thinking critically in that area. For example, standards for assessing the strength of inductive evidence for an empirical generalization cannot be sensibly or sensitively applied without knowing something about the nature of the phenomena covered by the generalization, including background theories concerning it and related phenomena. Similarly, thinking critically in deciding whether to accept or reject a moral judgement requires a clear understanding of the nature of the action or policy being judged, the context in which it is to be carried out, and the range of moral considerations relevant to the judgement. …*

Every area of intelligent human inquiry and practice, including science, art, law and morality, embodies within it practices of criticism by which proposed conclusions or ways of acting are tested, and previously accepted beliefs, practices and institutions are criticized and revised. Implicit in these practices are standards of critical assessment. It is these standards that critical thinkers must learn to use. They include not just rules of logic, but also standards of practical deliberation, standards of argumentation, standards used in developing plans of action, standards governing judgements made in the course of action (as in artistic and athletic performances), and standards governing inquiry and justification in specialized areas of study such as art, biology, history, literary criticism, mathematics, and technology. … Because verbal formulations of principles of critical thinking are abstracted from good critical practice, they typically do not tell a thinker all there is to know about the principles and how to apply them. For this reason they cannot be applied in a mechanical fashion. Rather, their abstractness gives them a vagueness that makes it necessary for the critical thinker to exercise judgement in interpreting them and determining what they require in any particular case. To acquire judgement of this sort, it is necessary to understand the practices of which the critical thinking principles are a part, and the point or purpose of these practices. It is also necessary to be acquainted with exemplars of the use of good judgement in applying the principles in a variety of contexts, because such examples provide the best indication of how the principle applies to particular cases. ...

Standards and principles of critical thinking are cultural artifacts that may be, and sometimes are, criticized and altered on the basis of our collective experience in using them. …

Although such standards are not readily revised, the appropriateness of any particular standard or the force it should have in a given context may always become a matter for critical reflection. …

One must also have certain commitments, attitudes or habits of mind that dispose him or her to use these resources to fulfill relevant standards and principles of good thinking. Passmore (1967: 197) aptly characterizes

---

the possession of these character traits as having a 'critical spirit'. Moreover, as Siegel (1988: 9) points out, the critical thinker's tendency to fulfill the standards and principles of good thinking cannot be mindless or simply the result of habituation. Rather, it must be based on a recognition of the value of critical thinking, i. e. its importance in fostering true belief and responsible action."

Bailin et al. 1999

[25] "*Thus, teaching critical thinking is best conceptualized not as a matter of teaching isolated abilities and dispositions, but rather as furthering the initiation of students into complex critical practices that embody value-commitments and require the sensitive use of a variety of intellectual resources in the exercise of good judgement. Initiation of children into these practices begins long before they enter school. By the time they are in primary school they are already making and criticizing judgements and arguments of various sorts, though their arguments and criticisms may not be very good. The educator's task is to continue the student's initiation in a more discriminating and self-conscious way, such that good critical practice is encouraged and poor practice is abandoned. This involves not simply teaching students standards and concepts of which they were previously ignorant, but also getting them to appreciate the value of changing some of their previously established commitments and practices. Although the long-range educational project is to develop competence in thinking critically in a variety of areas, the attainment of this end is necessarily a gradual process that can begin in the earliest years in school. Teaching students how to appraise evidential arguments in history or chemistry may have to await secondary school or university, but primary students can begin to learn important commitments and habits of mind, such as thinking reasons and truth are important, respecting others in discussion, being open-minded, and being willing to look at issues from others' points of view. They can learn a variety of critical concepts, such as those necessary for distinguishing between definitions and empirical statements; they can learn a number of heuristics, such as asking for examples when the meaning of a term is unclear and reminding themselves to double-check claims before accepting them as fact; and they can learn principles, such as trying to think of alternatives when deciding what is the best thing to do. As they become more mature they can expand the range of intellectual resources they are able and willing to employ and improve the judgement with which they employ them."*

Bailin et al. 1999

[26] *"Mystery meatloaf*
It's lunchtime. Five classmates have assembled in the cafeteria and are surveying the lunch choice. Present are Phil Gold, Nancy McGregor, Ravi Singh, Ahmed Ali, and Sophia Onassis.
Phil: Ah — mystery meatloaf — my favorite. I'll have a big piece with lots of gravy — and a double order of fries.
Gross!
Phil: Yeah— I guess you're right—double fries is a bit much.
Nancy: No— I mean — MEAT? RED meat, yuck! How can you eat that stuff?
Phil: So what's the big deal?
Haven't you heard ? Our Nancy's become a vegetarian.
Ahmed: No way! Not one of those granola-munching hippies?
Nancy: I just don't see how you can possibly bring yourself to eat another living creature. It's cruel… and inhumane!
Ravi: But animals eat other animals. It's just natural.
Ahmed: And besides, meat tastes so good. Just think of biting into a big, juicy, pink steak… mmmm.
Nancy: Ugh!
Phil: Everyone eats meat— at least all normal people. It's just some dumb cow.
Sophia: I've heard that animals used for meat are kept in horrible conditions.
Ravi: So now you believe everything you hear? That's not like you, Sophia. It's all a load of propaganda from those animal rights loonies.
Nancy: You guys are just a bunch of… cannibals!
Ahmed: Well, isn't she on her high horse, dictating to us what we should and shouldn't eat!
Nancy: Come on, Sophia — let's move to another table. I can't sit with the… insensitive boors!
Phil: Fine! Now at least we can eat in peace. Bring on the meatloaf…
You have likely come across scenes like this, possibly quite often. They are, unfortunately, fairly typical. There is a disagreement over an issue about which people feel strongly. The disagreement escalates into name-calling and high emotions and ends in misunderstanding.

But this conversation could go in a very different direction.

Mystery Meatloaf - Take II
Phil: Ah — mystery meatloaf — my favorite. I'll have a big piece with lots of gravy – and a double order of fries.
Nancy: I'll have the vegetarian lasagna please, with a side of yam fries.
Ahmed: What, no meatloaf today?
Ravi: Haven't you heard? Our Nancy's become a vegetarian.
Ahmed: No way! Why did you do something like that?
Nancy: It just finally got to me that I was eating an animal, another living creature, and that didn't seem right.
Ravi: But animals eat other animals. It's just natural.
Ahmed: And besides, it tastes so good.
Phil: Anyway, it's just a dumb cow, isn't it? It doesn't have thoughts or feelings like a person, does it?
Ravi: I'm pretty sure that animals feel pain. My dog sure howls when he gets his tail caught in the door.
Phil: Well, what about fish? They're not too with it.
Nancy: Some of my vegetarian friends do eat fish. I've been struggling with that one.
Sophia: I've heard that animals used for meat are kept in horrible conditions.
Ravi: I wonder if that's true or whether it's mostly propaganda from the animal rights folks?
Sophia: I haven't really checked it out ...
Ahmed: And there are some animals that live quite well. There are those free-range chickens who get to roam around and have lots of grain to eat and lead a normal chicken life (in fact, probably better than most). Until it's time to hop into the pot, that is. So is it OK to eat those free-range chicks?
Nancy: But it's still killing other living creatures for our selfish purposes. Why should we think that human beings have a right to do that?
Phil: It does bother me, though, when folks get so worked up about how we treat animals, especially cute ones with big eyes, and ignore all the people getting mistreated and even killed all over the world. Isn't that more important?
Sophia: Like the way all those movie stars and famous people protest about the seal hunt in Newfoundland but don't take any action about all the genocides happening around the world.
Ahmed: Wow – we've sure come up with a lot of questions. Though not many answers.
Sophia: I wonder ... maybe there's some way to go about trying to answer some of the questions. We couldn't be the first people to think about these issues. So we could have a look to see what ideas and information are out there.
Nancy: I'm sure there's information about the conditions in which animals are kept.
Ravi: And there must be research about whether different animals can feel pain, or even have other feelings.
Phil: And I'll bet other folks have thought about the moral issues about the treatment and rights of animals. I'd be interested in seeing what's been written about that.
Ravi: Though I don't expect that they'll all agree.
Phil: No, but that would at least give us some ideas to consider.
Ravi: And evaluate.
Sophia: I think it's worth a try. I don't know if we'll end up agreeing. Maybe. But even if we don't, at least we'll be able to think about the issues in a more informed way. And we'll be able to understand where the others are coming from.
Nancy: Now that would be progress!"

Bailin & Battersby 2016.

[27] "*Key aspects of critical thinking, as currently advocated by contemporary theorists, includes: 1. the claim that the notion is essentially normative in character; and 2. the claim that critical thinking involves two distinct components: a. skills or abilities of reason assessment and b. the dispositions to engage in and be guided by such assessments. ...*

So understood, critical thinking is a sort of good thinking. Therefore the notion of critical thinking is fundamentally a normative one, thus distinguishing this understanding of critical thinking from those, common in psychology, which treat the notion as descriptive, identifying particular psychological processes (Bailin et al

1999). To characterize a given episode of thinking as critical is to judge that it meets relevant standards or criteria of acceptability, and is thus appropriately thought of as good. …

Robert H. Ennis, for example, defines critical thinking as 'reasonable reflective thinking that is focused on deciding what to believe and do.' (Ennis, 1987), and offers a detailed list of abilities, skills and dispositions which thinking (and thinkers) must manifest if it is (they are) to qualify as critical. Matthew Lipman defines critical thinking as thinking that facilitates judgment because it relies on criteria, is self-correcting and is sensitive to context (Lipman 1991). Richard Paul analyses critical thinking in terms of the ability and the disposition to critically evaluate beliefs, their underlying assumptions, and the worldviews in which they are embedded (Paul 1990). Harvey Siegel characterizes the critical thinker as one who is 'appropriately moved by reasons' (Siegel 1988), and emphasizes the critical thinker's mastery of epistemic criteria which reasons must meet in order to be rightly judged to be good reasons, that is, reasons that justify beliefs, claims, judgments, and actions. Other authors, including John McPeck (1981, 1990) similarly emphasize the normative dimension of the concept. While these authors' accounts of critical thinking differ on many respects, and have their own emphases, they are nevertheless agreed on its essentially normative character (Bailin & Siegel 2003). … While some early treatments of critical thinking defined it only in terms of particular skills, … almost all recent philosophical discussion of it… regards critical thinking as involving both 1. skills or abilities of reason assessment and 2. a cluster of dispositions, habits of mind, and character traits, sometimes referred to collectively as the critical spirit (Siegel 1988)."

Siegel 2010

[28] "*The ideal calls for the fostering of certain skills and abilities, and for the fostering of a certain form of character. it is thus a general ideal of a certain sort of person, which sort of person it is the task of education to help to create. This aspects of the educational ideal of rationality aligns it with the complementary ideal of autonomy, since a rational person will – at least ideally – also be an autonomous one, capable of judging for himself/herself the justifiedness of candidate beliefs and the legitimacy of candidate values."*

Siegel 2010

[29] "*Ennis (1989) identifies a range of assumptions regarding domain specificity held by various theorists. For example, most researchers view background knowledge as a necessary but not sufficient condition for critical thinking. In addition, some researchers see the transfer of critical thinking skills across domains as unlikely unless students are provided with sufficient opportunities to practice these skills in a variety of domains and the students are explicitly taught to transfer. Finally, an even smaller number of researchers hold the view that general instruction in critical thinking skills is unlikely to be successful because critical thinking skills are inherently domain-specific (Ennis, 1989). Proponents of domain specificity include Willingham (2007), who argues that it is easier to learn to think critically within a given domain than it is to learn to think critically in a generic sense. Similarly, Bailin (2002) argues that domain-specific knowledge is necessary for critical thinking because what constitutes valid evidence, arguments, and standards tends to vary across domains. …*

Although McPeck (1990) concedes that there are a limited number of general thinking skills, he argues that the most useful thinking skills are those that are domain-specific. According to McPeck, the more general the thinking skill, the less helpful it is. …

Those who maintain that critical thinking skills and abilities are not domain-specific include Halpern (2001), who reviews evidence on the success of general instruction in critical thinking skills and concludes that such instruction has great potential. Lipman (1988) notes that critical thinking facilitates good judgment because it relies on criteria. These criteria may differ across domains, but the fundamental meaning of critical thinking remains the same. Van Gelder (2005, p. 43) argues that critical thinking is 'intrinsically general in nature,' which, paradoxically, is why critical thinking skills and abilities are notoriously difficult to transfer to new contexts."

Lai 2011

[30] "*Although the ability to think critically has always been important, it is a vital necessity for the citizens of the 21st century. …*

The workforce is one critical place where we can witness the dizzying pace of change. There is an increased demand for a new type ion worker – the knowledge worker or the symbol analyst – …. The information explosion *is yet another reason why we need to provide specific instruction in thinking. … Relevant, credible information has to be selected, interpreted, digested, evaluated, learned, and applied … The twin abilities of knowing how to learn and knowing how to think clearly about the rapidly proliferating information that we must select from are the most important intellectual skills of the 21st century. …*

For the first time in history, we have the ability to destroy all life on earth. The decisions that we make as individuals and as a society regarding the economy, conservation of the natural resources, and the development of nuclear weapons will affect future generations of all people around the world."

Halpern 2013

[31] "*Critical thinking refers to the use of cognitive skills or strategies that increase the probability of a desirable outcome. Critical thinking is purposeful, reasoned, and goal-directed. It is the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions. Critical thinkers use these skills appropriately, without prompting, and usually with conscious intent, in a variety of settings. That is, they are predisposed to think critically. When we think critically, we are evaluating the outcomes of our thought processes—how good a decision is or how well a problem is solved (Halpern, 1996, 1998). This definition is broad enough to encompass a variety of viewpoints, so critical thinking can be taught as argument analysis (see, for example, Kahane, 1997), problem solving (Mayer, 1992), decision-making (Dawes, 1988), or cognitive process (Rabinowitz, 1993). Regardless of the academic background of the instructor or the language used to describe critical thinking, all of these approaches share a set of common assumptions: there are identifiable critical thinking skills that can be taught and learned, and when students learn these skills and apply them appropriately, they become better thinkers."*

Halpern 1999

[32] "*What's 'critical' about critical thinking? The 'critical' part of critical thinking denotes an evaluation component. Sometimes the word 'critical' is used to convey something negative, as when we say 'she is a critical person'. However, evaluation can and should be a constructive reflection of positive and negative attributes. When we think critically, we are evaluating the outcomes of our thought processes – how good a decision is and how well a problem has been solved. Critical thinking also involves evaluating the thinking process – the reasoning that went into the conclusion we've arrived at or the kinds of factors considered in making g a decision."*

Halpern 2013

[33] "*Often there is noncritical, or more appropriately labelled, rote memorization or lower level thinking that is taught and tested in many classrooms, at all levels of education at the expense of higher ordered critical thinking. … Knowledge about a content area is critical to critical thinking ; no one can think critically about any topic without the necessary background information, but the facts alone are not enough. …*

Critical thinking skills are often referred to as 'higher order thinking skills' to differentiate them from simpler (i.e., lower order) thinking skills. Higher order skills are relatively complex; require judgment, analysis, and synthesis; and are to applied in rote and mechanical manner. Higher order thinking is thinking that is reflective, sensitive to the context, and self-monitored. Computational arithmetic, for example, is not a higher order skill, even though it is an important skill, because it involves the rote application of well-learned rules with little concern for context or other variables that would affect the outcome. By contrast, deciding which of two information sources is more credible is a higher order cognitive skill because it is a judgment task in which the variables that affect credibility are multidimensional and change with the context. In real life critical thinking skills are needed whenever we grapple with complex issues and messy, ill-defined problems."

Halpern 2007

[34] "*A general list of skills that would be applicable in almost any class would include understanding how cause is determined, recognizing and criticizing assumptions, analyzing means-goals relationships, giving reasons to support a conclusion, assessing degrees of likelihood and uncertainty, incorporating isolated data into a wider framework, and using an alogies to solve problems."*

Halpern 1998

---

[35] "*A short taxonomy of critical-thinking skills is proposed as a guide for instruction: (a) verbal reasoning skills--This category includes those skills needed to comprehend and defend against the persuasive techniques that are embedded in everyday language; (b) argument analysis skills--An argument is a set of statements with at least one conclusion and one reason that supports the conclusion. In real-life settings, arguments are complex, with reasons that run counter to the conclusion, stated and unstated assumptions, irrelevant information, and intermediate steps; (c) skills in thinking as hypothesis testing--The rationale for this category is that people function like intuitive scientists to explain, predict, and control events. These skills include generalizability, recognition of the need for an adequately large sample size, accurate assessment, and validity, among others; (d) likelihood and uncertainty--Because very few events in life can be known with certainty, the correct use of cumulative, exclusive, and contingent probabilities should play a critical role in almost every decision; (e) decision-making and problem-solving skills--In some sense, all of the critical-thinking skills are used to make decisions and solve problems, but the ones that are included here involve generating and selecting alternatives and judging among them. Creative thinking is subsumed under this category because of its importance in generating alternatives and restating problems and goals."*

Halpern 1998

[36] "*I am sometimes told that there is no such thing as critical thinking because different viewpoints are all a matter of opinion and that everyone has a right to his or her own opinion. They argue that a better way to think does not exist. I certainly agree that we all have the right to our own opinion; however some opinions are better than others. If, for example you believe that heavy alcohol consumption is good for pregnant women, you had better to back up this belief with a sound reasoning. …*

Similarly, everyone has the right to believe in phenomena such as astrology and extra-sensory perception, but there is no sound evidence to support the existence of these phenomena."

Halpern 2013

[37] "*Many authorities in higher education did not enthusiastically embrace the idea that college students should receive explicit instruction in how to think. Not that the academic community was opposed to good thinking, but many educators believed that it was a misguided effort. For example, Glaser (1984) cited abundant evidence of critical thinking failures in support of his argument that thinking skills are context-bound and do not transfer across academic domains. Glaser and other skeptics were partly correct. Better thinking is not a necessary outcome of traditional, discipline-based instruction. However, when thinking skills are explicitly taught for transfer, using multiple examples from several disciplines, students can learn to improve how they think in ways that transfer across academic domains."*

Halpern 1999

"We now have a considerable body of evidence that thinking skills courses and thinking skills instruction that is embedded in other courses can have positive effects that are transferable to many situations. …

The whole enterprise of learning how to improve thinking would be of little value if these skills were used only in the classroom or only on problems that are very similar to those presented in class. Ideally, critical thinking skills should be used to recognise and resist unrealistic campaign promises, circular reasoning, faulty probability estimates, weak arguments by analogy, or language designed to mislead whenever and wherever it is encountered. Critical thinkers should be able to solve or offer reasonable solutions to real-world problems, whether to solve the problem of nuclear war or how to set up a new video recorder."

Halpern 2013

[38] "*For example, students may be able to explain why correlation is not causation when presented with this question on an exam but still not recognize that this same principle is operating when they read that children who attend religious schools score higher on standardized tests than those who attend public schools. Specific instruction in recognizing the structure of correlational problems can improve the probability that students will recognize these problems, even when the topic is different."*

Halpern 1999

[39] "*It is important to separate the disposition or willingness to think critically from the ability to think critically. Some people may have excellent critical-thinking skills and may recognize when the skills are needed, but they*

also may choose not to engage in the effortful process of using them. This is the distinction between what people can do and what they actually do in real-world contexts. It is of no value to teach students the skills of critical thinking if they do not use them. Good instructional programs help learners decide when to make the necessary mental investment in critical thinking and when a problem or argument is not worth the effort. An extended session of generating alternatives and calculating probabilities is a reasonable response to a diagnosis of cancer; it is not worth the effort when the decision involves the selection of an ice-cream flavor."

Halpern 2013

[40] "*Critical thinking is effortful: it requires a concern for accuracy and the willingness to persist at difficult tasks Adeline suppress immediate Adeline easy responses. It requires an openness to new ideas, which some people find to be the most difficult component."*

Halpern 1999

[41] "*For example, making judgments about the likelihood of remembering a fact or event at some time in the future, or deciding how well a problem has been solved, or estimating one's own performance on a test of comprehension of complex prose. The underlying idea is that everyone needs to be able to assess how well they are thinking or how much they know ... about a topic to make reasoned decisions. Research has shown that when people have little knowledge of a content area (e.g. logical reasoning) they will misperceive their ability and take themselves to be much more higher in ability than they actually are …*

Students can become better thinkers and learners by developing the habit of monitoring their understanding and by judging the quality of their learning. …

Instruction to enhance the development of critical cal thinking skills should include a metacognitive component."

Halpern 1999

[42] "*Learning tasks, like real-world thinking tasks, should be rich in information. Some of the information available may not be relevant, and part of the learning exercise involves deciding which information is important to the problem. What is important in the teaching and learning of critical-thinking skills is what the learners are required to do with the information. Learning exercises should focus on the critical aspects of the problems and arguments that utilize the skills. The tasks should require thoughtful analysis and synthesis. For example, the repeated use of 'authentic' materials, or materials that are similar to real-world situations, is one teaching strategy to enhance transfer (Derry, Levin, & Schauble, 1995). Thinking skills need to be explicitly and consciously taught and then used with many types of examples so that the skill aspect and its appropriate use are clarified and emphasized."*

Halpern 1999

[43] "*1. Explicitly learn the skills of critical thinking*
2. Develop the dispositions for effortful thinking and learning
3. Direct learning activities in ways that increase the probability of trans contextual transfer (structure training)
4. Make metacognitive monitoring explicit and overt."

Halpern 2013

[44] "*Most instructional programs designed to teach critical thinking do not draw on contemporary empirical research in cognitive development as a potential resource. The developmental model of critical thinking outlined here derives from contemporary empirical research on directions and processes of intellectual development in children and adolescents. It identifies three forms of second-order cognition (meta-knowing) - metacognitive, metastrategic, and epistemological - that constitute an essential part of what develops cognitively to make critical thinking possible."*

Kuhn 1999

"... the developing cognitive competencies I describe as most relevant to critical thinking are metacognitive - rather than cognitive-competencies. In contrast to first-order cognitive skills that enable one to know about the world, metacognitive skills are second-order meta-knowing skills that entail knowing about one's own (and others') knowing. …

It should be noted, however, that a concept like metacognition, even if it does not go by this specific name, is by no means new to the philosophical literature on critical thinking. Indeed, something like metacognition figures in the definitions of critical thinking proposed by most educational philosophers who have addressed the topic. …

In contrast, an aspect of critical thinking that has received relatively little attention from educational philosophers is its developmental dimension. One important exception to this generalization, however, is the work of the educational philosopher who perhaps had the most to teach us about critical thinking - Dewey. Dewey, with his deep respect for and involvement in the empirical science of psychology, did not share the conception that prevails today of a dichotomy between the scientific study of human development and the practical world of children in schools. Instead, he saw schools as laboratories of human development, as experiments in the possibilities of human development in arranged environments. Repeatedly in his writings, Dewey made clear that the goal of education could not only be development (or what he called 'growth'). Education 'means supplying the conditions which foster growth' (Dewey, 1916, p. 56), *not toward a predetermined end but rather in the direction of 'an increase in the range and complexity of situations to which the child is capable of applying reasoned inquiry' (Cahan, 1994, p. 158). Dewey also made it clear that he saw the educator's task as a process of connecting with the young child's interests and purposes, but that one could not stop there. 'The real problem of intellectual education,' he said, 'is the transformation of more or less casual curiosity and sporadic suggestion into attitudes of alert, cautious, and thorough inquiry' (Dewey, 1933, p. 181)."*

Kuhn 1999

[45] "*A second distinctive characteristics of the present effort is that the developing cognitive competencies I describe as most relevant to critical thinking are metacognitive - rather than cognitive - competencies. In contrast to first-order cognitive skills that enable one to know about the world, metacognitive skills are second order meta-knowing skills that entail knowing about one's own (and others') knowing."*

Kuhn 1999

"The development of metacognitive understanding is essential to critical thinking because critical thinking by definition involves reflecting on what is known and how that knowledge is justified. Individuals with well-developed metacognitive skills are in control of their own beliefs in the sense of exercising conscious control over their evolution the face of external influences. They know what they think and ca justify why. Their skills in the conscious coordination of theory and evidence also put them in a position to evaluate the assertions of others."

Kuhn 1999

[46] "*Like pre-schoolers many older individuals blur the distinction between theory-based and evidence-based sources of their beliefs. Rather than seeing their theories as belief states subject to disconfirmation and representing theory and evidence as distinct entities to be reconciled with one another, they merge the two into a single representation of 'the way things are' with little apparent awareness of the source of their belief. Evidence serves merely to illustrate what one knows to be true…*

Theories may eventually change in response to discrepant evidence, but often with the individual manifesting little awareness or control of the process. Like young children… older participants in our studies are likely to deny that they ever held a belief different from the one they are now professing."

Kuhn 1999

[47] "*Each of the three kinds of meta-knowing that have been examined here-the metacognitive, metastrategic, and epistemological - is central to critical thinking. The development of metacognitive understanding is essential to critical thinking because critical thinking by definition involves reflecting on what is known and how that knowledge is justified. Individuals with well-developed metacognitive skills are in control of their own beliefs in the sense of exercising conscious control over their evolution in the face of external influences. They know what they think and can justify why. Their skills in the conscious coordination of theory and evidence also put them in a position to evaluate the assertions of others. Metastrategic skill is also essential to critical thinking. Those who have developed strong metastrategic skills apply consistent standards of evaluation across time and situations. They do not succumb to a view of a favored assertion as more probable than its alternatives because of its*

favored status and, therefore, subject to different standards of evolution. They also resist the temptation of local interpretation (Klahr, Fay, & Dunbar, 1993) of an isolated piece of evidence as supportive because it is considered out of the context of a broader pattern of which it is a part. The development of epistemological understanding may be the most fundamental underpinning of critical thinking. If knowledge is entirely objective, certain, and simply accumulates, unconnected to the human minds that do this knowing-as the absolutist conceives-or if knowledge is entirely subjective, subject only to the tastes and wishes of the knower-as the multiplist conceives-critical thinking and judgment are superfluous. People must see the point of thinking if they are to engage in it. Put simply by one of the multiplists in our studies, 'I feel it's not worth it to argue because everyone has their opinion.' In such cases, educators can undertake to teach intellectual skills, but the reasons to apply them will be missing."

Kuhn 1999

[48] "*In recent decades, philosophers, researchers and educators have focused their attention on a form of reasoning that combines reflection, justification, and the application and adaptation of individual reasoning; this has come to be known as critical thinking (Kuhn & Dean, 2004)."*

Kuhn 2018

[49] "*Each thirty-minute lesson was conducted in a classroom once every three days. The seminar included explicit instruction in regards to the development and analysis of arguments, counter-arguments and the value of trying to understand various views. It was designed to offer a nurturing environment and encouraged discussions regarding metacognitive practice. Students were asked to think about why they believe what they believe, and how they came to hold their beliefs. Discussions focused on understanding and analyzing the strength of reasons and arguments. The scrutiny of various claims and beliefs, along with an emphasis on open mindedness in the pursuit of good reasoning, were central themes of the seminar. Following the approach described by Kuhn and Udell (2003), the sessions included a great deal of dialogue and debate between groups of students who, through explicit instruction, had developed arguments for group discussion and analysis; together, students with opposing views discussed and analyzed their co-created arguments."*

Kuhn 2018

"The interview protocol contained six open-ended questions that focused on the following characteristics: judging the credibility of the source; identifying conclusions, reasons and assumptions (evidence); judging the quality of an argument (cogency); developing a position; open-mindedness (considering counter-arguments), and metacognitive awareness."

Kuhn 1992

[50] "*Twentieth-century psychologists have been pessimistic about teaching reasoning, prevailing opinion suggesting that people may possess only domain-specific rules, rather than abstract rules; this would mean that training a rule in one domain would not produce generalization to other domains. Alternatively, it was thought that people might possess abstract rules (such as logical ones) but that these are induced developmentally through self-discovery methods and cannot be trained. Research suggests a much more optimistic view: even brief formal training in inferential rules may enhance their use for reasoning about everyday life events. Previous theorists may have been mistaken about trainability, in part because they misidentified the kind of rules that people use naturally."*

Nisbett et al. 1987

[51] "*Our initial work on the use and trainability of inferential rules focused on a set of statistical rules that are derivable from the law of large numbers. We and our colleagues have found that people reason in accordance with the law of large numbers in a wide range of tasks and domains. For example, generalization often proceeds in accordance with the principle that larger samples are required when generalizing about populations that are more variable with respect to a given attribute than when generalizing about populations that are less variable. Three results support the view that people possess an abstract version of the law of large numbers and that improvements to it can transfer to a wide range of problem content. First, purely abstract rule training produced improvement in both the frequency and the quality of statistical answers. Second, the abstract rule training effect was substantial across all three problem domains: training improved statistical reasoning for problems that people rarely think of in terms of probability just as much as it did for problems that people almost always think*

of in probabilistic terms. Since highly abstract statistical rules can be taught in such a way that they can be applied to a great range of everyday life events, is the same true of the even more abstract rules of deductive logic? We can report no evidence indicating that this is true, and we can provide some evidence indicating that it is not."

Nisbett et al. 1987

[52] "*Inferences frequently violate principles of statistics, economics, logic, and basic scientific methodology."*

Nisbett 2015

[53] "*Initially I was quite dubious that a course or two dealing with one or another approach to reasoning could have the kind of impact on people that long exposure to the concepts had on me. The XXth century skepticism about the possibility of teaching reasoning continued to influence my thinking. I could not have been more mistaken. It turns out that the courses people take in college do affect inferences about the world - often very markedly. Rules of logic, statistical principles such as how large numbers and regression to the mean, principles of scientific methodology such as how to establish control groups when making assertions about cause and effect, classical economic principles, and decision theory concepts all influence the way people think about problems that crop up in everyday life. The affect how people reason about athletic events, what procedures they think are best for going about the process of hiring someone, and event their approach to such minor questions as whether they should finish a meal that isn't very tasty. ...*

Most impressively, we sometimes questioned subjects weeks later, in contexts where they didn't know they were being studied, such as telephone polls allegedly being conducted by survey researchers. We were delighted to find that people often retained substantial ability to apply the concepts to ordinary problems outside the laboratory contexts in which the concepts had been taught."

Nisbett 2015

[54] "*Remember that all perceptions, judgments, and beliefs are inferences and are not direct readouts of reality. This recognition should prompt an appropriate humility about just how certain we should be about our judgments, as well as recognition that the views of other people that differ from our own may have more validity than our intuitions tell us they do.*
Remember that incidental, irrelevant perceptions and cognitions can affect our judgment and behavior.
Be alert of the possible role of heuristics in producing judgments.
Pay more attention to the context.
These injunctions can become part of the mental equipment you use to understand the world."

Nisbett 2015

[55] "*We don't normally think of forming impressions of individual's personality as a statistical process consisting of sampling a population of events, but they are exactly that. and framing them in that way makes us both more cautious about some kinds of personality ascriptions and better able to predict individual's behavior in the future. ...*

(1) frame everyday life events in such a way that the relevance of statistical principles is obvious and you can make contact with them, and (2) code the events in such a way that approximate versions of statistical rules can be applied to them… Once you have the knack of framing real-world problems as statistical ones and coding their elements in such a way that statistical heuristics can be applied, those principles seem to pop up magically to help you solve a given problem - often without conscious awareness that you're applying a rough-and-ready version of a statistical principle."

Nisbett 2015

"What counts as an explanation (for everything from why your friend acts in such an annoying way to why a product launch failed)? How can we tell the difference between events that are causally related and events that are merely associated to one another in time or place? What kinds of knowledge can be considered certain and what kinds only conjectural? What are the characteristics of a good theory - in science and in everyday life? How can we tell the difference between theories that can be falsified and theories that cannot?"

Nisbett 2015

[56] *"The distinction between two kinds of thinking, one fast and intuitive, the other slow and deliberative, is both ancient in origin and widespread in philosophical and psychological writing. Such a distinction has been made by many authors in many fields, often in ignorance of the related writing of others (Frankish & Evans, 2009). Our particular interest is in dual-process accounts of human reasoning and related higher cognitive processes, such as judgment and decision-making. Such theories have their origins in the 1970s and 1980s (Evans, 1989; Wason & Evans, 1975) and have become the focus of much interest in contemporary research on these topics (Barbey & Sloman, 2007; Evans, 2007a, 2008; Evans & Over, 1996; Kahneman, 2011; Kahneman & Frederick, 2002; S. A. Sloman, 1996; Stanovich, 1999, 2011; Stanovich & West, 2000). Over a similar period, dual-process theories have proved popular in the psychology of learning (e.g., Dienes & Perner, 1999; Reber, 1993; Sun, Slusarz, & Terry, 2005) and especially in social cognition, which has the greatest proliferation of dual-processing labels and theories (see Chaiken & Trope, 1999; Epstein, 1994; Kruglanski & Orehek, 2007; Smith & DeCoster, 2000). Originally, dual-process theories in these different fields developed independently, although there have more recently been attempts to connect them. One consequence has been the development of broad dual-system theories that attempt to link a wide range of attributes to two systems of thought that are believed to underlie intuitive and reflective processing, respectively (Epstein, 1994; Evans & Over, 1996; Reber, 1993; Stanovich, 1999). Following Stanovich (1999), these are often referred to as Systems 1 and 2. As the popularity of dual-process and dual-system theories has increased, so too have the voices of criticism, as illustrated in the opening quotations. Critics have pointed to the multitude of dual-processing accounts, the vagueness of their definition, and the lack of coherence and consistency in the proposed cluster of attributes for two-system accounts. They have questioned the evidence on which such claims are made and have argued that single-process accounts can explain the data (Gigerenzer & Regier, 1996; Keren & Schul, 2009; Kruglanski et al., 2003; Kruglanski & Gigerenzer, 2011; Osman, 2004). Here we collaborate for the first time to respond to these various critiques. It is important that we do so, as although a number of these criticisms have some force to them (and have been acknowledged in our own recent writing), we believe that the dual-processing distinction is nonetheless strongly supported by a wide range of converging experimental, psychometric, and neuroscientific methods. In general, these critiques are problematic because they attack not any particular theory but rather a class of theories, effectively treating all dual-process and dual- system theories alike. However, all dual-process theories are not, by any means, the same. Our own work has developed dual-process theories of reasoning and decision-making, but even in this domain, there is much in the writings of other authors with which we have disagreements."*

Evans & Stanovich 2013

[57] "*Critical thinking is highly valued in educational writings if not in practice. Despite a substantial literature on the subject, for many years the area of critical thinking was notorious for its conceptual difficulties. For example, years ago Cuban (1984) lamented that 'defining thinking skills, reasoning, critical thought, and problem solving is troublesome to both social scientists and practitioners. Troublesome is a polite word; the area is a conceptual swamp' (p. 676). There has been some progress in elucidating the concept of critical thinking since the time of Cuban's statement, but we shall argue here that educational theory is on the verge of an even more stunning conceptual advance in the area of critical thinking. Education is beginning to understand the critical thinking concept by relating it to the constructs of intelligence and rational thought. In fact, modern cognitive science provides a coherent framework for understanding the relation between critical thinking, intelligence, and rational thought."*

Evans & Stanovich 2010

[58] "*Cognitive scientists recognize two types of rationality: instrumental and epistemic. The simplest definition of instrumental rationality is behaving in the world so that you get exactly what you most want, given the resources (physical and mental) available to you. Somewhat more technically, we could characterize instrumental rationality as the optimization of the individual's goal fulfillment. Economists and cognitive scientists have refined the notion of optimization of goal fulfillment into the technical notion of expected utility. The model of rational judgment used by decision scientists is one in which a person chooses options based on which option has the largest expected utility (see Baron, 2008; Dawes, 1998; Hastie & Dawes, 2001; Wu, Zhang, & Gonzalez, 2004). The other aspect of rationality studied by cognitive scientists is termed epistemic rationality. This aspect of rationality concerns how well beliefs map onto the actual structure of the world. Epistemic rationality is sometimes called theoretical rationality or evidential rationality (see Audi, 1993, 2001; Foley, 1987; Harman, 1995; Manktelow, 2004; Over, 2004). Instrumental and epistemic rationality are related. The aspect of beliefs*

that enter into instrumental calculations (that is, tacit calculations) are the probabilities of states of affairs in the world."

Evans & Stanovich 2010

[59] "*In the critical thinking literature, the ability to evaluate evidence and arguments independently of one's prior beliefs and opinions is a skill that is strongly emphasized (Baron, 2008; Dole & Sinatra, 1998; Ennis, 1987, 1996; Kuhn, 2005; Lipman, 1991; Paul, 1984, 1987; Ritchhart & Perkins, 2005; Sternberg, 1997, 2001, 2003; Wade & Tavris, 1993). The disposition toward such unbiased reasoning is almost universally viewed as a characteristic of good thinking. …*

Neimark (1987) emphasizes how associations built up over time will tend to activate a decision for us automatically and unconsciously if we are not reflective and cannot detach from situational cues. The danger of response patterns that are determined too strongly by overlearned cues is a repeated theme in the heuristics and biases literature of cognitive science (e.g., Arkes, 1991; Evans, 2003, 2006, 2007, 2008; Kahneman, 2003; Kahneman & Frederick, 2002; Stanovich, 2003, 2004, 2009; Wilson & Brekke, 1994). ... Many tasks in the heuristics and biases branch of the reasoning literature might be said to involve some type of decontextualization skill (Kahneman, 2003; Stanovich, 2003). Tasks are designed to see whether reasoning processes can operate independently of interfering context (world knowledge, prior opinion, vivid examples)."

Evans & Stanovich 2010

[60] "*Degrees of rationality can be assessed in terms of the number and severity of such cognitive biases that individuals display. Failure to display a bias becomes a measure of rational thought."*

Stanovic & Stanovich 2010

[61] "*Virtually all attempts to classify heuristics and biases tasks end up utilizing a dual-process framework because most of the tasks in the literature on heuristics and biases were deliberately designed to pit a heuristically triggered response against a normative response generated by the analytic system. As Kahneman (2000) notes, 'Tversky and I always thought of the heuristics and biases approach as a two-process theory' (p. 682). Since Kahneman and Tversky launched the heuristics and biases approach in the 1970s, a wealth of evidence has accumulated in support of the dual-process approach. Evidence from cognitive neuroscience and cognitive psychology converges on the conclusion that mental functioning can be characterized by two different types of cognition having somewhat different functions and different strengths and weaknesses (Brainerd & Reyna, 2001; Evans, 2003, 2008, 2009; Evans & Over, 1996, 2004; Feldman Barrett, Tugade, & Engle, 2004; Greene, Nystrom, Engell, Darley, & Cohen, 2004; Kahneman & Frederick, 2002, 2005; Lieberman, 2003; McClure, Laibson, Loewenstein, & Cohen, 2004; Metcalfe & Mischel, 1999; Sloman, 1996, 2002; Stanovich, 1999, 2004). There are many such theories (over 20 dual-process theories are presented in a table in Stanovich, 2004) and they have some subtle differences, but they are similar in that all distinguish autonomous from non-autonomous processing. The two types of processing were termed systems in earlier writings, but theorists have been moving toward more atheoretical characterizations; we therefore follow Evans (2009) in using the terms type 1 and type 2 processing. The defining feature of type 1 processing is its autonomy. Type 1 processes are termed autonomous because (a) their execution is rapid, (b) their execution is mandatory when the triggering stimuli are encountered, (c) they do not put a heavy load on central processing capacity (i.e., they do not require conscious attention), (d) they do not depend on input from high-level control systems, and (e) they can operate in parallel without interfering with each other or with type 2 processing. Type 1 processing would include behavioral regulation by the emotions, the encapsulated modules for solving specific adaptive problems that have been posited by evolutionary psychologists, processes of implicit learning, and the automatic firing of overlearned associations (see Evans, 2007, 2008; Stanovich, 2004, 2009). Type 2 processing contrasts with type 1 processing on each of the critical properties that define the latter. Type 2 processing is relatively slow and computationally expensive—it is the focus of our awareness. Many type 1 processes can operate at once in parallel, but only one (or a very few) type 2 thoughts can be executing at once—type 2 processing is thus serial processing. Type 2 processing is often language-based. One of the most critical functions of type 2 processing is to override type 1 processing. All of the different kinds of type 1 processing (processes of emotional regulation, Darwinian modules, associative and implicit learning processes) can produce responses that are irrational in a particular context if not overridden. In order to override type 1 processing, type 2 processing must display at least two (possibly related) capabilities. One is the capability of interrupting type 1 processing and suppressing its response tendencies. Type 2 processing thus involves inhibitory mechanisms of the type that have been the*

focus of recent work on executive functioning (Hasher, Lustig, & Zacks, 2007; Miyake, Fried- man, Emerson, & Witzki, 2000; Zelazo, 2004).”

Stanovic & Stanovich 2010

[62] “*This concern for the efficiency of information processing as opposed to its rationality is mirrored in the status of intelligence tests. They are measures of efficiency but not rationality—a point made clear by considering a distinction that is very old in the field of psychometrics. Psychometricians have long distinguished typical performance situations from optimal (sometimes termed maximal) performance situations (see Ackerman, 1994, 1996; Ackerman & Heggestad, 1997; Ackerman & Kanfer, 2004; see also, Cronbach, 1949; Matthews, Zeidner, & Roberts, 2002). Typical performance situations are unconstrained in that no overt instructions to maximize performance are given and the task interpretation is determined to some extent by the participant. The goals to be pursued in the task are left somewhat open. The issue is what a person would typically do in such a situation, given few constraints. Typical performance measures are measures of the reflective mind—they assess in part goal prioritization and epistemic regulation. In contrast, optimal performance situations are those where the task interpretation is determined externally. The person performing the task is instructed to maximize performance and is told how to do so. Thus, optimal performance measures examine questions of efficiency of goal pursuit—they capture the processing efficiency of the algorithmic mind. All tests of intelligence or cognitive aptitude are optimal performance assessments, whereas measures of critical or rational thinking are often assessed under typical performance conditions. The difference between the algorithmic mind and the reflective mind is captured in another well-established distinction in the measurement of individual differences—the distinction between cognitive ability and thinking dispositions. The former are, as just mentioned, measures of the efficiency of the algorithmic mind. The latter travel under a variety of names in psychology—thinking dispositions or cognitive styles being the two most popular. Many thinking dispositions concern beliefs, belief structure and, importantly, attitudes toward forming and changing beliefs. Other thinking dispositions that have been identified concern a person’s goals and goal hierarchy. Examples of some thinking dispositions that have been investigated by psychologists are: actively open-minded thinking, need for cognition (the tendency to think a lot), consideration of future consequences, need for closure, superstitious thinking, and dogmatism (Cacioppo, Petty, & Feinstein 1996; Kruglanski & Webster, 1996; Norris & Ennis, 1989; Schommer-Aikins, 2004; Stanovich, 1999, 2009; Sternberg, 2003; Sternberg & Grigorenko, 1997; Strathman, Gleicher, Boninger, & Scott Edwards, 1994). The literature on these types of thinking dispositions is vast and our purpose is not to review that literature here. It is only necessary to note that the types of cognitive propensities that these thinking disposition measures reflect are the tendency to collect information before making up one’s mind, to seek various points of view before coming to a conclusion, to think extensively about a problem before responding, to calibrate the degree of strength of one’s opinion to the degree of evidence available, to think about future consequences before taking action, to explicitly weigh pluses and minuses of situations before making a decision, and to seek nuance and avoid absolutism. In short, individual differences in thinking dispositions include assessing variation in people’s goal management, epistemic values, and epistemic self-regulation—differences in the operation of reflective mind. They are all psychological characteristics that underpin rational thought and action.”*

Stanovic & Stanovich 2010

[63] “*To be rational, a person must have well-calibrated beliefs and must act appropriately on those beliefs to achieve goals—both properties of the reflective mind. The person must, of course, have the algorithmic-level machinery that enables him or her to carry out the actions and to process the environment in a way that allows the correct beliefs to be fixed and the correct actions to be taken. Thus individual differences in rational thought and action can arise because of individual differences in intelligence (the algorithmic mind) or because of individual differences in thinking dispositions (the reflective mind). …*

To think rationally means adopting appropriate goals, taking the appropriate action given one’s goals and beliefs, and holding beliefs that are commensurate with available evidence. Standard intelligence tests do not assess such functions (Perkins, 1995, 2002; Stanovich, 2002, 2009; Sternberg, 2003, 2006). For example, although intelligence tests do assess the ability to focus on an immediate goal in the face of distraction, they do not assess at all whether a person has the tendency to develop goals that are rational in the first place. Likewise, intelligence tests provide good measures of how well a person can hold beliefs in short-term memory and manipulate those beliefs, but they do not assess at all whether a person has the tendency to form beliefs rationally when presented with evidence. And again, similarly, intelligence tests give good measures of how *efficiently a person processes information that has been provided, but they do not at all assess whether the person is a critical assessor of information as it is gathered in the natural environment. …*

Substantial empirical evidence indicates that individual differences in thinking dispositions and intelligence are far from perfectly correlated. Many different studies involving thousands of subjects (e.g., Ackerman & Heggestad, 1997; Austin & Deary, 2002; Baron, 1982; Bates & Shieles, 2003; Cacioppo et al., 1996; Eysenck, 1994; Goff & Ackerman, 1992; Kanazawa, 2004; Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Zeidner & Matthews, 2000) have indicated that measures of intelligence display only moderate to weak correlations (usually less than .30) with some thinking dispositions (e.g., actively open-minded thinking, need for cognition) and near zero correlations with others (e.g., conscientiousness, curiosity, diligence).

Other important evidence supports the conceptual distinction made here between algorithmic cognitive capacity and thinking dispositions. For example, across a variety of tasks from the heuristics and biases literature, it has consistently found that rational thinking dispositions will predict variance in these tasks after the effects of general intelligence have been controlled (Bruine de Bruin, Parker, & Fischhoff, 2007; Klaczynski, Gordon, & Fauth, 1997; Klaczynski & Lavallee, 2005; Klaczynski & Robinson, 2000; Kokis et al., 2002; Macpherson & Stanovich, 2007; Newstead, Handley, Harley, Wright, & Farrelly, 2004; Parker & Fischhoff, 2005; Sá & Stanovich, 2001; Stanovich & West, 1997, 1998c, 2000; Toplak, Liu, Macpherson, Toneatto, & Stanovich, 2007; Toplak & Stanovich, 2002).

Measures of thinking dispositions tell us about the individual's goals and epistemic values—and they index broad tendencies of pragmatic and epistemic self-regulation at a high level of cognitive control. The empirical studies cited indicate that these different types of cognitive predictors are tapping separable variance, and the reason that this is to be expected is because cognitive capacity measures such as intelligence and thinking dispositions map on to different levels in the tripartite model."

Stanovic & Stanovich 2010

[64] "*A simple example of miserly processing is discussed by Kahneman and Frederick (2002). They describe a simple experiment in which people were asked to consider the following puzzle: 'A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost?'*

Many people offer the response that first comes to mind—10¢– without thinking further and realizing that this cannot be right. The bat would then have to cost $1.10 and the total cost would be $1.20 rather than the required $1.10. People often do not think deeply enough to make this simple correction though, and many students at very selective universities will answer incorrectly and move on to the next problem before realizing that their shallow processing has led them to make an error. Frederick (2005) has found that large numbers of highly selected students at MIT, Princeton, and Harvard, when given this and other similar problems, are cognitive misers like the rest of us. The correlation between intelligence and a set of similar items is quite modest, in the range of .40 to .50 (Gilhooly & Murphy, 2005).

Many other biases of the cognitive miser show correlations no greater than those shown in the Frederick bat-and-ball problem. In fact, some cognitive biases are almost totally dissociated from intelligence. Myside bias, for example, is virtually independent of intelligence (Macpherson & Stanovich, 2007; Sá, Kelley, Ho, & Stanovich, 2005; Stanovich & West, 2007, 2008a, 2008b; Toplak & Stanovich, 2003). Individuals with higher IQs in a university sample are no less likely to process information from an egocentric perspective than are individuals with relatively lower IQs.

Irrational behavior can occur not just because of miserly processing tendencies but also because the right mindware (cognitive rules, strategies, knowledge, and belief systems) is not available to use in decision-making."

Stanovic & Stanovich 2010

[65] "*Metacomponents are higher order executive processes used to plan what one is going to do, monitor it while one is doing it, and evaluate it after it is done. The metacomponents include recognising that a problem exists, defining the nature of the problem, deciding on a set of steps for solving the problem, ordering these steps into a coherent strategy, deciding upon a form of mental representation for information, allocating one's time and resources in solving a problem, monitoring one's solution to a problem as the problem is being solved, and utilising feedback regarding problem solving after one's problem solving has been completed*."

Sternberg 1986

[66] "*Consider, for example, the performance components of induction. These include encoding stimuli, mapping relations between relations, applying relations from one domain to another, justifying potential responses, and responding.*"

Sternberg 1986

[67] "*Three such composants are selective encoding, which involves screening relevant from irrelevant information, selective combination, which involves putting together the relevant information in a coherent and organized way; and selective comparison, which involves relating old, previously known information to new, about to be learned information.*"

Sternberg 1986

[68] "*The lessons are : acquiring expertise in critical thinking is hard ; practice in critical thinking skills themselves enhances skills ; the transfer of skills must be practiced ; some theoretical knowledge is required ; diagramming arguments (argument mapping) promotes skills ; and students are prone to belief preservation.*"

Gelder 2005

[69] "*Humans are not naturally critical. Indeed, like ballet, critical thinking is a highly contrived activity. Running is natural; nightclub dancing is less so; but ballet is something people can only do well with many years of painful, expensive, dedicated training. Evolution did not intend us to walk on the ends of our toes, and whatever Aristotle might have said, we were not designed to be all that critical either. Evolution does not waste effort making things better than they need to be, and homo sapiens evolved to be just logical enough to survive, while competitors such as Neanderthals and mastodons died out.*"

Gelder 2005

[70] "*From the cognitive scientist's point of view, the mental activities that are typically called critical thinking are actually a subset of three types of thinking: reasoning, making judgments and decisions, and problem solving. I say that critical thinking is a subset of these because we think in these ways all the time, but only sometimes in a critical way. Deciding to read this article, for example, is not critical thinking. But carefully weighing the evidence it presents in order to decide whether or not to believe what it says is. Critical reasoning, decision-making, and problem solving—which, for brevity's sake, I will refer to as critical thinking—have three key features: effectiveness, novelty, and self-direction. Critical thinking is effective in that it avoids common pitfalls, such as seeing only one side of an issue, discounting new evidence that disconfirms your ideas, reasoning from passion rather than logic, failing to support statements with evidence, and so on. Critical thinking is novel in that you don't simply remember a solution or situation that is similar enough to guide you. For example, solving a complex but familiar physics problem by applying a multi-step algorithm isn't critical thinking because you are really drawing on memory to solve the problem. But devising a new algorithm is critical thinking. Critical thinking is self-directed in that the thinker must be calling the shots: We wouldn't give a student much credit for critical thinking if the teacher were prompting each step he took.*"

Willingham 2007

[71] "*Can critical thinking actually be taught? Decades of cognitive research point to a disappointing answer: not really. People who have sought to teach critical thinking have assumed that it is a skill, like riding a bicycle, and that, like other skills, once you learn it you can apply it in any situation. Research from cognitive science shows that thinking is not that sort of skill. The processes of thinking are intertwined with the content of thought (that is, domain knowledge).*"

Willingham 2007

[72] "*… if you remind a student to 'look at an issue from multiple perspectives' often enough, he will learn that he ought to do so, but if he doesn't know much about an issue, he can't think about it from multiple perspectives. You can teach students maxims about how they ought to think, but without background knowledge and practice, they probably will not be able to implement the advice they memorize.*"

Willingham 2007

[73] "*Studies of the Philosophy for Children program may be taken as typical. Two researchers identified eight studies that evaluated academic outcomes and met minimal research-design criteria. (Of these eight, only one had been subjected to peer review.) Still, they concluded that three of the eight had identifiable problems that clouded the researchers' conclusions. Among the remaining five studies, three measured reading ability, and one of these reported a significant gain. Three studies measured reasoning ability, and two reported significant gains. And, two studies took more impressionistic measures of student's participation in class (e.g., generating ideas, providing reasons), and both reported a positive effect*."

Willingham 2007

[74] "*Despite the difficulties and general lack of rigor in evaluation, most researchers reviewing the literature conclude that some critical thinking programs do have some positive effect. But these reviewers offer two important caveats. First, as with almost any educational endeavor, the success of the program depends on the skill of the teacher. Second, thinking programs look good when the outcome measure is quite similar to the material in the program. As one tests for transfer to more and more dissimilar material, the apparent effectiveness of the program rapidly drop*s."

Willingham 2007

[75] "*If knowledge of how to solve a problem never transferred to problems with new surface structures, schooling would be inefficient or even futile—but of course, such transfer does occur. When and why is complex, but two factors are especially relevant for educators: familiarity with a problem's deep structure and the knowledge that one should look for a deep structure*."

W Willingham 2007

"Here's an example: A treasure hunter is going to explore a cave up on a hill near a beach. He suspected there might be many paths inside the cave so he was afraid he might get lost. Obviously, he did not have a map of the cave; all he had with him were some common items such as a flashlight and a bag. What could he do to make sure he did not get lost trying to get back out of the cave later? The solution is to carry some sand with you in the bag, and leave a trail as you go, so you can trace your path back when you're ready to leave the cave. About 75 percent of American college students thought of this solution—but only 25 percent of Chinese students solved it.6 The experimenters suggested that Americans solved it because most grew up hearing the story of Hansel and Gretel, which includes the idea of leaving a trail as you travel to an unknown place in order to find your way back. The experimenters also gave subjects another puzzle based on a common Chinese folk tale, and the percentage of solvers from each culture reversed. (To read the puzzle based on the Chinese folk tale, and the tale itself, go to www.aft.org/ pubs-reports/american_educator/index.htm.) It takes a good deal of practice with a problem type before students know it well enough to immediately recognize its deep structure, irrespective of the surface structure, as Americans did for the Hansel and Gretel problem."

Willingham 2007

"They are little chunks of knowledge—like 'look for a problem's deep structure' or 'consider both sides of an issue'—that students can learn and then use to steer their thoughts in more productive directions." « Le problème de CT genre de stratégie métacognitive est qu'elle ne peut pas porter beaucoup plus loin qu'à amener à se répéter au bon moment la stratégie métacognitive. » "Thus, a student who has been encouraged many times to see both sides of an issue, for example, is probably more likely to spontaneously think 'I should look at both sides of this issue' when working on a problem. … Unfortunately, metacognitive strategies can only take you so far. Although they suggest what you ought to do, they don't provide the knowledge necessary to implement the strategy."

Willingham 2007

"Understanding and using conditional probabilities is essential to scientific thinking because it is so important in reasoning about what causes what. But people's success in thinking this way depends on the particulars of how the question is presented. Studies show that adults sometimes use conditional probabilities successfully, but fail to do so with many problems that call for it. Even trained scientists are open to pitfalls in reasoning about conditional probabilities (as well as other types of reasoning). Physicians are known to discount or misinterpret

new patient data that conflict with a diagnosis they have in mind, and Ph.D.-level scientists are prey to faulty reasoning when faced with a problem embedded in an unfamiliar context."

Willingham 2007

[76] "*But critical thinking is very different. As we saw in the discussion of conditional probabilities, people can engage in some types of critical thinking without training, but even with extensive training, they will sometimes fail to think critically. This understanding that critical thinking is not a skill is vital*."

Willingham 2007

[77] "*It tells us that teaching students to think critically probably lies in small part in showing them new ways of thinking, and in large part in enabling them to deploy the right type of thinking at the right time…*

What do all these studies boil down to? First, critical thinking (as well as scientific thinking and other domain-based thinking) is not a skill. There is not a set of critical thinking skills that can be acquired and deployed regardless of context. Second, there are metacognitive strategies that, once learned, make critical thinking more likely. Third, the ability to think critically (to actually do what the metacognitive strategies call for) depends on domain knowledge and practice."

Willingham, D. T. (2007). Critical thinking: Why it is so hard to teach? *American federation of teachers summer 2007, p. 8-19*.

[78] "*For example, in one experiment researchers taught college students principles for evaluating evidence in psychology studies—principles like the difference between correlational research and true experiments, and the difference between anecdote and formal research (Bensley & Spero, 2014). These principles were incorporated into regular instruction in a psychology class, and their application was practiced in that context. Compared to a control group that learned principles of memory, students who learned the critical thinking principles performed better on a test that required evaluation of psychology evidence.*

There is even evidence that critical thinking skills can be taught and applied in complex situations under time pressure. In one experiment, officers in the Royal Netherlands Navy received training in the analysis of complex battlefield problems in a high- fidelity tactical simulator. They were first taught a sequence of steps to undertake when analyzing this sort of problem, and then underwent a total of 8 hours of training on surface warfare problems, with feedback from an expert. The critical outcome measure was performance (without feedback) in a new surface warfare problem, as well as performance on air warfare problems. Judges assessed the quality of participant's action contingency plans, and those receiving the training outperformed control subjects (Helsdingen et al., 2010).

When we think of critical thinking, we think of something bigger than its domain of training. When I teach students how to evaluate the argument in a set of newspaper editorials, I am hoping that they will learn to evaluate arguments generally, not just those they read, and not just those they would find in other editorials. This aspect of critical thinking is called transfer, and the research literature evaluating how well critical thinking skills transfer to new problems is decidedly mixed. …

It is not useful to think of critical thinking skills, once acquired, as broadly applicable. Wanting students to be able to 'analyse, synthesise and evaluate' information sounds like a reasonable goal. But analysis, synthesis, and evaluation mean different things in different disciplines. Literary criticism has its own internal logic, its norms for what constitutes good evidence and a valid argument. These norms differ from those found in mathematics, for example. And indeed, different domains—science and history, say—have different definitions of what it means to 'know' something. Thus, our goals for student critical thinking must be domain-specific. An overarching principle like 'think logically' is not a useful goal. …

Experimental evidence shows that an expert does not think as well outside her area of expertise, even in a closely related domain. She is still better than a novice, but her skills do not transfer completely. For example, knowledge of medicine transfers poorly among subspecialties; neurologists do not diagnose cardiac cases well (Rikers, Schmidt, & Boshuizen, 2002). Expertise in writing is similarly encapsulated; a technical writer who specialises in writing instruction pamphlets for home electronics can't write newspaper articles (Kellogg, 2018). Perhaps most surprisingly the analytic abilities of professional philosophers do not extend to everyday judgments. Philosophers are no less susceptible than average adults to being swayed by irrelevant features of problems like question order or wording (Schwitzgebel & Cushman, 2015)."

Willingham 2019

[79] "*First, identify what is meant by critical thinking in each domain. Be specific. What tasks showing critical thinking should a high school graduate be able to do in mathematics, history, and other subjects? It is not useful to set a goal that students 'think like historians,' or 'learn the controversies surrounding historical events.' If students are to read as historians do, they need to learn specific skills like interpreting documents in light of their sources, corroborating them, and putting them in historical context. Notably, skillful reading is different in other disciplines. Scientists believe that the source of a document is irrelevant so long as it is trustworthy. And unlike historical documents, scientific documents are written in a consistent format. Learning to read like a scientist means, in part, learning the conventions of this format. These skills should be explicitly taught and practiced—there is evidence that simple exposure to this sort of work without explicit instruction is less effective (Abrami et al., 2008; Halpern, 1998; Heijltjes, Van Gog, & Paas, 2014). In addition, it is clear that educators will have to pick and choose which skills their students will learn. Even across the long thirteen years of schooling, time is limited. Second, identify the domain content that students must know. We have seen that domain knowledge is a crucial driver of thinking skills. For example, sourcing historical documents means interpreting their content in light of the author, the intended audience, and circumstances under which the author wrote. It is not enough to know that a letter was written by an army sergeant to his wife just before the Battle of Romani. The student must know enough about the historical context to understand how this sourcing information ought to influence his or her interpretation of the letter. Fourth, educators must decide which skills should be revisited across years. Studies show that even if content is learned quite well over the course of half of a school year, about half will be forgotten in three years (Pawl et al., 2012). That doesn't mean there's no value in exposing students to content just once; most students will forget much but they'll remember something, and for some students, an interest may be kindled. But when considering skills we hope will stick with students for the long term, we should plan on at least three to five years of practice (Bahrick, 1984; Bahrick & Hall, 1991). Most of the time, this practice will look different—it will be embedded in new skills and content. But this revisiting should be assured and planned.*"

Willingham 2019

[80] *Another area of disagreement among critical thinking researchers is the extent to which critical thinking skills and abilities can be transferred to new contexts. For example, researchers have noted that students may exhibit critical thinking skills and abilities in one context, or domain, but fail to do so in another (Willingham, 2007). This issue is closely related to that of the domain-specificity of critical thinking. For example, those maintaining that critical thinking is completely domain-specific are more likely to be skeptical of students' abilities to transfer critical thinking skills from one domain to another (Ennis, 1989). Accepted wisdom within cognitive psychology holds that spontaneous transfer to new contexts is rare (Kennedy et al., 1991; Pithers & Soden, 2000; Willingham, 2007). Others, however, are more sanguine about the possibility of student transfer, particularly if students are given opportunities to practice critical thinking skills in multiple domains and contexts and if students are taught specifically to transfer those skills (Kennedy et al., 1991). McPeck (1990), a staunch proponent of domain specificity, notes that his approach does not preclude the transfer of critical thinking skills and abilities to real-world contexts, particularly when instruction emphasizes authentic learning activities that represent problems encountered in daily life.*

Lai 2011

[81] *Existing published assessments of critical thinking are numerous, and include the California Critical Thinking Skills Test (Facione, 1990), the Cornell Critical Thinking Tests (Ennis & Millman, 2005), the Ennis-Weir Critical Thinking Essay Test (Ennis & Weir, 1985), and the Watson-Glaser Critical Thinking Appraisal (Watson & Glaser, 1980). As Ku (2009) points out, these instruments vary widely in both purpose and item format. However, as Kennedy et al. (1991) note, none of these tests are intended for use with students below the fourth-grade level. Moreover, these assessments tend to be general critical thinking assessments rather than subject-specific.*

Lai 2011

[82] *Despite differences among the three schools of thought and their approaches to defining critical thinking, there exist areas for agreement. First, researchers of critical thinking typically agree on the specific abilities encompassed by the definition, which include:*

- *analyzing arguments, claims, or evidence (Ennis, 1985; Facione, 1990; Halpern, 1998; Paul, 1992);*

- *making inferences using inductive or deductive reasoning (Ennis, 1985; Facione, 1990; Paul, 1992; Willingham, 2007);*
- *judging or evaluating (Case, 2005; Ennis, 1985; Facione, 1990; Lipman, 1988; Tindal & Nolet, 1995); and*
- *making decisions or solving problems (Ennis, 1985; Halpern, 1998; Willingham, 2007).*

Other abilities or behaviors identified as relevant to critical thinking include asking and answering questions for clarification (Ennis, 1985); defining terms (Ennis, 1985); identifying assumptions (Ennis, 1985; Paul, 1992); interpreting and explaining (Facione, 1990); reasoning verbally, especially in relation to concepts of likelihood and uncertainty (Halpern, 1998); predicting (Tindal & Nolet, 1995); and seeing both sides of an issue (Willingham, 2007).

Lai 2011

[83] *Some researchers have argued that the link between critical thinking and metacognition is self-regulation. For example, the APA Delphi report includes self-regulation as one component skill of critical thinking (Facione, 1990). Schraw et al. (2006) draw connections between metacognition, critical thinking, and motivation under the umbrella of self-regulated learning, which they define as "our ability to understand and control our learning environments" (p. 111). Self-regulated learning, in turn, is seen as comprising three components: cognition, metacognition, and motivation. The cognitive component includes critical thinking, which Schraw and associates explain consists of identifying and analyzing sources and drawing conclusions. However, others have argued that critical thinking and metacognition are distinct constructs. For example, Lipman (1988) has pointed out that metacognition is not necessarily critical, because one can think about one's thought in an unreflective manner. McPeck, on the other hand, argues that the ability to recognize when a particular skill is relevant and to deploy that skill is not properly a part of critical thinking but actually represents general intelligence (1990). At the very least, metacognition can be seen as a supporting condition for critical thinking, in that monitoring the quality of one's thought makes it more likely that one will engage in high-quality thinking.*

Lai 2011

[84] *Many researchers working in the area of critical thinking lament the poor state of critical thinking in most educated adults and children. For example, Halpern (1998) points to research from the field of psychology, concluding that many, if not most, adults fail to think critically in many situations. Kennedy et al., (1991) and Van Gelder (2005) have likewise concluded that many adults lack basic reasoning skills. Halpern (1998) cites the example that large numbers of people profess to believe in paranormal phenomena, despite a lack of evidence in support of such things. Halpern attributes such failures not to the inability to reason well but to simple "bugs" in reasoning. She argues that human beings are programmed to look for patterns, particularly in the form of cause-and-effect relationships, even when none exist. Van Gelder (2005) echoes this sentiment, characterizing humans as "pattern-seekers and story-tellers" (p. 42). This inclination results in a tendency to jump to the first explanation that makes intuitive sense without carefully scrutinizing alternative possibilities, a phenomenon that Perkins, Allen, & Hafner (1983) have termed "makes-sense epistemology" (p. 286). Moreover, the general public often finds "personal experience" to be more compelling evidence than a carefully conducted, scientific study. Given these natural tendencies toward deficient reasoning, Halpern warns that we should not expect to see dramatic improvements in critical thinking over time as a result of instructional interventions. Improvements in critical thinking, when they do occur, are slow and incremental (Halpern, 1998).*

Lai 2011

[85] *According to Kuhn's (1999) theoretical framework, metacognitive knowing characterizes the first stirrings of critical thought in very young children. There are two distinct stages within metacognitive knowing. The first stage is called Realism and is typically achieved between the ages of 3 and 5. This stage is characterized by the belief that assertions are expressions of someone's belief, and as such, may depart from reality. Thus, the child is able to identify true and false statements. Prior to reaching this stage, children regard beliefs and assertions as isomorphic with reality. "In other words, the world is a simple one in which things happen and we can tell about them. There are no inaccurate renderings of events" (p. 19).*

According to Kuhn's framework (1999), the second stage of metacognitive knowing, typically achieved by 6 years of age, allows the child to be aware of sources of knowledge and further, to distinguish between theory *and evidence. In other words, prior to reaching this second stage, the child has difficulty distinguishing evidence for the claim that an event has occurred from the causal theory that makes occurrence of the event plausible. In other words, is something true because it makes intuitive sense or because there is empirical evidence for it? Kuhn describes a study (Kuhn & Pearsall, 1998) in which children were shown a series of pictures depicting two runners competing in a race. The last picture shows one of the runners holding up a trophy and smiling. When children were asked who won the race, most children correctly indicated that the runner represented in the final photo was the winner. However, when asked to justify this claim, younger children tended to cite causal theories ("because he is wearing fast shoes") rather than evidence in support of the claim ("because he is holding a trophy"). According to Kuhn, by the second stage of metacognitive knowing children are able to make this distinction.*

Based on the empirical research in meta-memory, Kuhn's framework (1999) also portrays meta-strategic knowing in two stages. According to Kuhn, during the first stage, typically achieved during middle childhood, children begin to understand the value of cognitive strategies in aiding cognition. A child who has reached this stage recognizes that a memory strategy such as categorization will aid recall and tends to effectively manage and deploy cognitive resources during problem solving (Kuhn, 1999). The second stage of meta-strategic knowing may not be achieved at all. If it is attained, it is typically reached during adolescence and adulthood. According to Kuhn, this stage is characterized by consistent and appropriate strategy selection from a repertoire of available strategies. Thus, the individual monitors strategy.

Based on the empirical research in meta-memory, Kuhn's framework (1999) also portrays meta-strategic knowing in two stages. According to Kuhn, during the first stage, typically achieved during middle childhood, children begin to understand the value of cognitive strategies in aiding cognition. A child who has reached this stage recognizes that a memory strategy such as categorization will aid recall and tends to effectively manage and deploy cognitive resources during problem solving (Kuhn, 1999). The second stage of meta-strategic knowing may not be achieved at all. If it is attained, it is typically reached during adolescence and adulthood. According to Kuhn, this stage is characterized by consistent and appropriate strategy selection from a repertoire of available strategies. Thus, the individual monitors strategy According to Kuhn (1999), the second stage in epistemological understanding, labeled the Multiplist Epistemological position, tends to be prevalent during adolescence. During this stage, the individual acknowledges that experts can disagree and actually relinquishes the idea of certainty. A person in this stage moves to the opposite end of the subjectivity-objectivity continuum, vis-à-vis those in the Absolutist stance. Instead of viewing the world as inherently and objectively knowable, individuals in this stage perceive the world as a completely subjective place. In other words, "because all people have a right to their opinions, all opinions are equally right" (p. 22). Kuhn points out that many people become permanently stuck in this phase.

Finally, Kuhn (1999) argues that the last stage in epistemological understanding (and critical thinking), to which only a minority of people will ever progress, is known as Epistemological Metaknowing. According to Kuhn's framework (1999), at this stage the individual is able to balance the subjective and objective, recognizing a multiplicity of valid This person uses judgment, evaluation, and argumentation to sift through opinions and arrive at those that are most valid. Not all opinions are valued equally; rather, reason, logic, and empirical evidence can be used to privilege certain positions over others (Kuhn, 1999). This person uses judgment, evaluation, and argumentation to sift through opinions and arrive at those that are most valid. Not all opinions are valued equally; rather, reason, logic, and empirical evidence can be used to privilege certain positions over others (Kuhn, 1999).

Lai 2011

[86] *Reasoning is generally seen as a means to improve knowledge and make better decisions. However, much evidence shows that reasoning often leads to epistemic distortions and poor decisions. This suggests that the function of reasoning should be rethought. Our hypothesis is that the function of reasoning is argumentative. It is to devise and evaluate arguments intended to persuade. Reasoning so conceived is adaptive given the exceptional dependence of humans on communication and their vulnerability to misinformation. A wide range of evidence in the psychology of reasoning and decision-making can be reinterpreted and better explained in the light of this hypothesis. Poor performance in standard reasoning tasks is explained by the lack of argumentative context. When the same problems are placed in a proper argumentative setting, people turn out to be skilled arguers. Skilled arguers, however, are not after the truth but after arguments supporting their views. This explains the notorious confirmation bias. This bias is apparent not only when people are actually arguing, but*

also when they are reasoning proactively from the perspective of having to defend their opinions. Reasoning so motivated can distort evaluations and attitudes and allow erroneous beliefs to persist. Proactively used reasoning also favors decisions that are easy to justify but not necessarily better. In all these instances traditionally described as failures or flaws, reasoning does exactly what can be expected of an argumentative device: Look for arguments that support a given conclusion, and, ceteris paribus, favor conclusions for which arguments can be found.

Mercier & Sperber 2011

[87] *We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based. CT is essential as a tool of inquiry. As such, CT is a liberating force in education and a powerful resource in one's personal and civic life. While not synonymous with good thinking, CT is a pervasive and self-rectifying human phenomenon. The* ***ideal critical thinker i****s habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair- minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit. Thus, educating good critical thinkers means working toward this ideal. It combines developing CT skills with nurturing those dispositions which consistently yield useful insights and which are the basis of a rational and democratic society.*

Facione, 1990

[88] "*Communication brings vital benefits, but carries a major risk for the audience of being accidentally or intentionally misinformed. Nor is there any failsafe way of calibrating one's trust in communicated information so as to weed out all and only the misinformation. Given that the stakes are so high, it is plausible that there has been ongoing selective pressure in favour of any available cost-effective means to least approximate such sorting. Since there are a variety of considerations relevant to the granting or withholding of epistemic trust, we will explore the possibility that different abilities for epistemic vigilance may have emerged in biological and cultural evolution, each specialising in a particular kind of relevant considerations. Factors affecting the acceptance or rejection of a piece of communicated information may have to do either with the source of the information—who to believe; or with its content—what to believe."*

Sperber et al 2010

[89] *"We claim that humans have a suite of cognitive mechanisms for epistemic vigilance, targeted at the risk of being misinformed by others. Here we present this claim and consider some of the ways in which epistemic vigilance works in mental and social life. Our aim is to integrate into a coherent topic for further research a wide range of assumptions developed elsewhere by ourselves or others, rather than to present detailed arguments for each." ...*

How reliable are others as sources of information? In general, they are mistaken no more often than we are—after all, 'we' and 'they' refer to the same people—and they know things that we don't know. So it should be advantageous to rely even blindly on the competence of others. Would it be more advantageous to modulate our trust by exercising some degree of vigilance towards the competence of others? That would depend on the cost and reliability of such vigilance. But in any case, the major problem posed by communicated information has to do not with the competence of others, but with their interests and their honesty. While the interests of others often overlap with our own, they rarely coincide with ours exactly. In a variety of situations, their interests are best served by misleading or deceiving us. It is because of the risk of deception that epistemic vigilance may be not merely advantageous but indispensable if communication itself is to remain advantageous. …

People stand to gain immensely from communication with others, but this leaves them open to the risk of being accidentally or intentionally misinformed, which may reduce, cancel, or even reverse these gains. The fact that communication is so pervasive despite this risk suggests that people are able to calibrate their trust well enough to make it advantageous on average to both communicator and audience (Sperber, 2001; Bergstrom et al., 2006). For this to happen, the abilities for overt intentional communication and epistemic vigilance must have evolved together, and must also develop together and be put to use together.

Sperber et al 2010

---

90 *"Some contents are intrinsically believable even if they come from an untrustworthy source. Examples include tautologies, logical proofs, truisms, and contents whose truth is sufficiently evidenced by the act of communication itself (e.g. saying, 'Je suis capable de dire quelques mots en français'). Other contents are intrinsically unbelievable even if they come from a trustworthy source. Examples include logical contradictions, blatant falsehoods, and contents whose falsity is sufficiently evidenced by the act of communication itself (e.g. saying, 'I am mute'). ... checking takes place against the narrow context of beliefs used in the search for a relevant interpretation of the utterance."*

Sperber et al 2010

91 *"Judgements about the trustworthiness of informants may be more or less general or contextualised. You may think, 'Mary is a trustworthy person', meaning it both epistemically and morally, and therefore expecting what Mary says to be true, what she does to be good, and so on. Or you may trust (or mistrust) someone on a particular topic in specific circumstances: 'You can generally trust Joan on Japanese prints, but less so when she is selling one herself'. Trust can be allocated in both these ways, but how do they compare from a normative point of view ?*

A reliable informant must meet two conditions: she must be competent, and she must be benevolent. That is, she must possess genuine information (as opposed to misinformation or no information), and she must intend to share that genuine information with her audience (as opposed to making assertions she does not regard as true, through either indifference or malevolence). Clearly, the same informant may be competent on one topic but not on others, and benevolent towards one audience in certain circumstances, but not to another audience or in other circumstances. This suggests that trust should be allocated to informants depending on the topic, the audience, and the circumstances. However such precise calibration of trust is costly in cognitive terms, and, while people are often willing to pay the price, they also commonly rely on less costly general impressions of competence, benevolence and overall trustworthiness."

Sperber et al 2010

92 *There is a growing body of research on the development of children's epistemic vigilance (for reviews, see e.g. Koenig and Harris, 2007; Heyman, 2008; Clement, in press; Corriveau and Harris, in press; Nurmsoo et al., in press). This shows that even at a very early age, children do not treat all communicated information as equally reliable. At 16 months, they notice when a familiar word is inappropriately used (Koenig and Echols, 2003). By the age of two, they often attempt to contradict and correct assertions that they believe to be false (e.g. Pea, 1982). These studies challenge the widespread assumption that young children are simply gullible*

Sperber et al 2010

93 *"One group of 3-, 4- and 5-year olds watched and listened as two speakers narrated a short passage from the story of 'Curious George'. One spoke English with a native (North-American) accent. The other spoke English with a foreign (Spanish) accent. A second group of children of the same age watched and listened as the two speakers narrated a short passage from 'Jabberwocky'—the nonsense poem by Lewis Carroll. Although syntactically well-formed, the sentences in this passage were not meaningful so that any differences in trust following this induction could not be attributed to differential comprehension of the two speakers. Following both types of induction, children were given an opportunity to seek and endorse information about the use of four unfamiliar artefacts from the two speakers. They offered conflicting demonstrations of how to use any given artefact. For example, one speaker looked through a plastic sprinkler attachment as if it were a telescope, whereas the other speaker held it to her mouth and blew in it. Children preferred to seek and endorse information from the native-accented speaker. This preference was equally strong in all three age groups and equally strong following the meaningful, 'Curious George' induction and the meaningless, 'Jaberwocky' induction. Note that the induction phase and the test phase of this experiment differed in both modality and domain. The induction phase involved audible differences in accent. The test phase involved visible differences in tool use. Nevertheless, children used the audible cues of group membership to guide their learning about tool use."*

Harris & Corriveau 2011

94 *"...suppose that children encounter two informants who make conflicting claims that are novel and therefore impossible for children to adjudicate themselves. However, the claims made by one informant elicit approval*

from bystanders, whereas the claims made by the other elicit disapproval. Do children use such bystanders' reactions to moderate their trust in the novel claims made by each informant? To examine this possibility, we had 4-year olds watch as two speakers produced conflicting names for a series of unfamiliar objects. For example, faced with the sprinkler attachment, one speaker might call it a 'feppin' and the other might call it a 'merval'. The two bystanders reacted differently to the two speakers. Having listened to one, they nodded and smiled. Having listened to the other, they shook their head and frowned. Subsequently, children were asked for their judgement. They were reminded that one speaker had called it a feppin and the other had called it a merval—what did they think? Children overwhelmingly endorsed the speaker who had attracted bystanders' approval rather than disapproval. In the next stage of the experiment, we tested if children would continue to regard the speaker who had received bystanders' approval as more trustworthy even in the absence of any feedback from the bystanders. To assess this possibility, the two bystanders left the room, and testing continued as before with the two informants making conflicting claims about unfamiliar objects. Children continued to display selective trust in the two speakers—they were more likely to endorse the names supplied by the speaker who had received bystanders' approval even though, at this point in the experiment, the bystanders were no longer present and could supply no cues. By implication, the cultural typicality of the two speakers—the extent to which their claims had met with approval versus disapproval—led children to regard one of them as a more trustworthy informant. However, an alternative interpretation of these results is that children did not conclude that the two informants differed in terms of cultural typicality but in terms of likeability. After all, in expressing their approval, the bystanders had smiled at one informant, and in expressing their disapproval, they had frowned at the other. Arguably, children preferred to endorse the speaker whom they inferred to be more likeable, as indexed by the bystanders' reactions. In a follow-up study, we again had two informants as well as an additional pair of adults who sided with one informant and not with the other. However, we altered the way in which this endorsement was expressed. Several unfamiliar objects were set out on a table and the experimenter asked the adults to say which of them was, for example, 'a slod'. Three of the adults pointed to the same object, whereas the fourth—the lone dissenter—pointed to a different object. This pattern was repeated for four trials with the same person always in the role of a lone dissenter. After watching the adults' responses, children were invited to express their view. As in the previous study, children strongly favoured the majority view, effectively endorsed by three of the adults, as opposed to the minority view endorsed by only a single adult.

Harris & Corriveau 2011

[95] *"No matter how non-selective children are in what they learn from others, they are selective in whom they learn from. We have identified two broad classes of heuristics—one class helps children to select among informants with whom they have had previous interactions, and the second class helps children to differentiate among relatively unfamiliar informants whom they have just met...*

Within the first class, children display two biases. First, they display a preference for the information supplied by a familiar caregiver versus a stranger (provided that they have not developed an avoidant relationship with that caregiver). Second, children prefer information supplied by someone who has proven to be a reliable source of information in the past. Taken together, these two biases are likely to converge on a proclivity for vertical cultural learning—a bias to endorse and imitate the claims and demonstrations of adults who have a record of providing reliable care, accurate information, or both. ... The second class of biases enables children to differentiate among informants with whom they have had no protracted interaction. As noted, this class leads children to prefer informants who appear to be culturally typical, either in the sense that the informants signal that they belong to the same group as the children (because of the way that they speak or look) or in the sense that other potential informants assent to, rather than dissent from, the information offered by the informant. These biases are likely to promote oblique and horizontal cultural learning that is relatively conservative. When children encounter someone who is not a familiar caregiver, they will be more inclined to accept guidance from that person if he or she appears to belong to, and receives endorsement from, the children's cultural group. Stated differently, children's receptivity to both oblique and horizontal learning does not extend to all-comers. They are less likely to trust information that is provided by members of another cultural group or by deviants from within their own group."

Harris & Corriveau 2011

[96] *Epistemic vigilance directed at informants yields a variety of epistemic attitudes (acceptance, doubt or rejection, for instance) to the contents communicated by these informants. There is some evidence that three-year-old children are aware of attitudes such as endorsement or doubt (Fusaro and Harris, 2008), and are also*

aware that assertions can be stronger or weaker (Sabbagh and Baldwin, 2001; Birch et al., 2008; Matsui et al., 2009). Children are able to make sense of comments on the reliability of what is communicated (e.g. Fusaro and Harris, 2008, Clement et al., 2004). As a result, they can take advantage of the epistemic judgments of others, and enrich their own epistemological understanding and capacity for epistemic vigilance in doing so.

Sperber et al 2010

[97] "*Likewise, the acquisition of other metacognitive and metastrategic skills is a gradual process. Early strategies for coordinating theory and evidence are replaced with better ones, but there is not a stage-like change from using an older strategy to a newer one. … However, metastrategic competence does not appear to routinely develop in the absence of instructions*

Morris et al. 2012

[98] *"A reliable informant must meet two conditions: she must be competent, and she must be benevolent. That is, she must possess genuine information (as opposed to misinformation or no information), and she must intend to share that genuine information with her audience (as opposed to making assertions she does not regard as true, through either indifference or malevolence). Clearly, the same informant may be competent on one topic but not on others, and benevolent towards one audience in certain circumstances, but not to another audience or in other circumstances. This suggests that trust should be allocated to informants depending on the topic, the audience, and the circumstances. However such precise calibration of trust is costly in cognitive terms, and, while people are often willing to pay the price, they also commonly rely on less costly general impressions of competence, benevolence and overall trustworthiness."*

Sperber et al 2010

[99] *A striking illustration of the tendency to form general judgments of trustworthiness on the basis of very limited evidence is provided in a study by Willis and Todorov (2006). Participants were shown pictures of faces, for either a mere 100 milliseconds or with no time limit, and asked to evaluate the person's trustworthiness, competence, likeability, aggressiveness and attractiveness. Contrary to the authors' expectations, the correlation between judgments with and without time limit was not greater for attractiveness —which is, after all, a property of a person's appearance—than for trustworthiness, while the correlations for aggressiveness and competence were a relatively low. One might wonder if such split-second judgments of trustworthiness have any basis at all, but what this experiment strongly suggests is that looking for signs of trustworthiness is one of the first things we do when we see a new face (see also Ybarra et al., 2001). There is a considerable social psychology literature suggesting that people's behaviour is determined to a significant extent not by their character but by the situation (Ross and Nisbett, 1991; Gilbert and Malone, 1995). If so, judging that someone is generally trustworthy may be a case of the 'fundamental attribution error' (Ross, 1977): that is, the tendency, in explaining or predicting someone's behaviour, to overestimate the role of psychological dispositions and underestimate situational factors. But even without appealing to character psychology, it is possible to defend the view that some people are more generally trustworthy than others, and are to some extent recognisable as such."*

Sperber et al 2010

[100] *"The very social success which is almost a defining feature of cultural information might suggest that (except in cases of cultural conflict) it is uncritically accepted. We will argue, however, that here too epistemic vigilance is at work, but that it needs appropriate cultural and institutional development to meet some of the epistemic challenges presented by cultural information... Often, information spreads through a group from a single source, and is accepted by people along the chains of transmission because they trust the source rather than because of any evidence or arguments for the content. If so, the crucial consideration should be the trustworthiness of the original source. If each person who passes on the information has good independent reasons for trusting the source, this should give people further along the chain good reasons for also trusting the source, and thus for accepting the content originally conveyed. However, people's reasons for trusting the source are in general no more independent of one another than their reasons for accepting the content...*

It might seem, then, that people are simply willing, or even eager, to accept culturally transmitted information without exercising ordinary epistemic vigilance towards it. Boyd, Richerson and Henrich have argued that there is an evolved conformist bias in favour of adopting the behaviour and attitudes of the majority of members of one's community (e.g. Boyd and Richerson, 1985; Henrich and Boyd, 1998). Csibra and Gergely (2009) have *argued that people in general, and children in particular, are eager to acquire cultural information, and that this may bias them towards interpreting (and even over-interpreting) communicated information as having cultural relevance, and also towards accepting it. An alternative (or perhaps complementary) hypothesis is that people do exercise some degree of epistemic vigilance towards all communicated information, whether local or cultural, but that their vigilance is directed primarily at information originating in face to face interaction, and not at information propagated on a larger scale. For instance, people may be disposed to pay attention to the problems raised by the non-independence of testimonies, or by discrepancies in their contents, when they are blatantly obvious, as they often are when they occur in face to face interaction, but not otherwise. On a population scale, these problems can remain unnoticed although, on reflection, they are likely to be pervasive. All kinds of beliefs widely shared in the community may propagate throughout a culture by appealing to individual trust in converging testimonies. The trust is not blind, but the epistemic vigilance which should buttress it is short-sighted...*

So far, the picture we have sketched of epistemic vigilance on a population scale is somewhat grim. Mechanisms for epistemic vigilance are not geared to filtering information transmitted on such a large scale. Even if we are right to claim that these mechanisms exist, they do not prevent mistaken ideas, undeserved reputations and empty creeds from invading whole populations. However, we did note that it is important not to jump from the fact that people are seriously, even passionately, committed to certain ideas, and expect others to be similarly committed, to the conclusion that the commitment involved is clearly epistemic. It may be that the content of the ideas matters less to you than who you share them with, since they may help define group identities. When what matters is the sharing, it may be that contents which are unproblematically open to epistemic evaluation would raise objections within the relevant social group, or would be too easily shared beyond that group. So, semi-propositional contents which can be unproblematically accepted by just the relevant group may have a cultural success which is negatively correlated with their epistemic value."

Sperber et al 2010

[101] *"In a number of domains, there are institutional procedures for evaluating the competence of individuals, making these evaluations public through some form of certification, and sanctioning false claims to being so certified. Medical doctors, professors, judges, surveyors, accountants, priests, and so on are generally believed to be experts in their field because they have shown strong evidence of their expertise to experts who are even more qualified. Of course, these procedures may be inadequate or corrupt, and the domain may itself be riddled with errors; but still, such procedures provide clear and easily accessible evidence of an individual's expertise… The institutional organisation of epistemic vigilance is nowhere more obvious than in the sciences, where observational or theoretical claims are critically assessed via social processes such as laboratory discussion, workshops, conferences, and peer review in journals. The reliability of a journal is itself assessed through rankings, and so on (Goldman, 1999). Social mechanisms for vigilance towards the source and vigilance towards the content interact in many ways. In judicial proceedings, for instance, the reputation of the witness is scrutinised in order to strengthen or weaken her testimony. In the sciences, peer review is meant to be purely content-oriented, but is influenced all too often by the authors' prior reputation (although blind reviewing is supposed to suppress this influence), and the outcome of the reviewing process in turn affects the authors' reputation. Certification of expertise, as in the granting of a PhD, generally involves multiple complex assessments from teachers and examiners, who engage in discussion with the candidate and among themselves; these assessments are compiled by educational institutions which eventually deliver a reputation label, 'PhD', for public consumption."*

Sperber et al 2010

[102] "*Our main aim in doing so is to suggest that, to a significant extent, these social mechanisms are articulations of psychological mechanisms linked through extended chains of communication, and, in some cases through institutional patterning (Sperber, 1996). In these population scale articulations, psychological mechanisms combine with cognitive artefacts (e.g. measuring instruments), techniques (e.g. statistical tests of confidence), and procedures (e.g. for cross-examination) to yield distributed epistemic assessment systems (Heintz, 2006) which should be seen as a special kind of distributed cognitive system (Hutchins, 1996)."*

Sperber et al 2010

[103] *"The form of gullibility that is the main target here has the three following traits. First, it views gullibility as widespread: people would be very often fooled into accepting empirically unfounded messages. Second, it views*

gullibility as often applying to costly beliefs, beliefs that lead to painful rituals, expensive purchases, risky rebellions, or harmful complacence. Third, it views gullibility as being mostly source-based: stemming from the undue influence of focal sources, often authority figures, be they religious leaders, demagogues, TV anchors, celebrities, and so forth. Most accusations of gullibility reviewed above share these traits. I will refer to this view of gullibility as strong gullibility."

Mercier 2017

[104] *"Inconsistencies between background beliefs and novel information easily lead to belief updating. If John sees a green elephant in the yard, he updates his beliefs accordingly. John can afford to do this because his perceptual and inferential mechanisms do not attempt to mislead him. By contrast, in the case of communicated information, the honesty of the sender is open to question. This means that communicated information that is inconsistent with a receiver's background beliefs should be, on average, less likely to lead to belief revision than similar information obtained through perception (in the absence of contrary evidence provided by trust or arguments, see below, and Sperber et al., 2010). We would thus rely on plausibility checking, a mechanism that detects inconsistencies between background beliefs and communicated information, and that tends to reject communicated information when such inconsistencies emerge. ...*

There is substantial evidence that people detect inconsistencies between their background beliefs and communicated information, that such inconsistencies tend to lead to reject communicated information, and that information that is more inconsistent with one's prior beliefs is more likely to be rejected. ...

For instance, in a typical advice taking experiment, participants have to form an opinion about a given question—typically a numerical opinion, such as 'how much does the individual in this picture weighs?' They are then confronted with the opinion of another participant, and they can revise their opinion on this basis. When no relevant factor, such as expertise, differentiates the participant receiving the advice from the participant giving the advice, the advice tends to be heavily discounted in favor of the participant's own opinion (e.g., Yaniv & Kleinberger, 2000). Moreover, the advice is more discounted when it is further away from the participant's own opinion (Yaniv, 2004)."

Mercier 2017

[105] *"A major motivation for seeking advice is the need to improve judgment accuracy and the expectation that advice will help. An abundance of studies have shown that combining multiple sources of information improves estimation in the long run, in a variety of domains ranging from perceptual judgment to business forecasting (e.g., Armstrong, 2001; Sorkin, Hayes, & West, 2001; Yaniv, 1997). Aside from accuracy, there are also social reasons for seeking advice, which we consider only briefly here. Accountants performing complex audit tasks tend to solicit advice for self-presentational reasons and to increase the justification for their decisions (Kennedy, Kleinmuntz, & Peecher, 1997). Indeed, seeking advice implies sharing with others the responsibility for the outcome of a decision (Harvey & Fischer, 1997). One might argue, however, that even self-presentational reasons for seeking advice are rooted in the belief on the part of the individual or the organization that consulting someone else's opinion could improve one's final decision… A basic dilemma in using advice involves the amount of weight to place on others' opinions. Receiving advice often exposes decision makers to a potential conflict between their initial opinions and the advice. Consider a manager who believes that a certain new product is likely to gain success and is thus worthy of further development. The manager then receives a lukewarm expert opinion of her idea. How might she revise her opinion? The key question in many practical situations is to decide just how much weight ought to be assigned to a particular piece of advice. In particular, a decision maker's weighting policy might entail completely ignoring the other opinion, some adjustment of one's own opinion towards the other, or complete adoption of the other opinion... Previous work on the use of advice in decision-making suggests a self/other effect whereby individuals tend to discount advice and favor their own opinion. In a judgmental estimation task (Yaniv & Kleinberger, 2000) respondents formed a final opinion on the basis of their initial opinion and a piece of advice. Rather than using equal weighting, respondents tended to place a higher weight on their own opinion than on the advisor's opinion. Even though the decision makers were sensitive to the quality of the advice (good vs poor), they tended to discount both good and poor advice. In a cue-learning study by Harvey and Fischer (1997), respondents shifted their estimates about 20–30% towards the advisor's estimates. Lim and O'Connor (1995) found that, in combining their prior personal forecasts and advisory (statistical) forecasts, judges weighted their own forecasts more heavily than the statistical forecasts."*

Yaniv 2004

106 *"Receivers use a wide variety of cues to infer senders' trustworthiness. Some cues relate to the competence of the sender. A competent sender is a sender who is likely to have formed reliable beliefs. Cues to competence can be traits of senders such as dispositions (intelligence, diligence), or acquired expertise (being skilled in mechanics, say). Cues to competence can also be local, such as differences of perceptual access (who witnessed the crime).*

Other cues relate to the sender's benevolence. A benevolent sender is a sender who is likely to send messages that positively take into account the receivers' interests (Barber, 1983). Thus, benevolence entails more than the absence of lying. If a sender sends a message that only benefits the sender, and not the receiver, she would not be deemed benevolent, even if the message is not an outright lie. For instance, a friend who recommends a restaurant on the basis of preferences she knows not to be shared by her audience would not be very benevolent. Like cues to competence, cues to benevolence can be traits, stable features of senders that make them more likely to be benevolent toward the relevant receiver (relatedness to the receiver, say). Cues to benevolence can also be local. In particular, attention should be paid to how the interests of the sender are served by the acceptation of the message being evaluated. Self-interested messages should arouse suspicion.

...receivers take cues to trustworthiness into account in an overall sensible way, and that they lower their trust in senders who were committed to messages that proved unreliable… In line with this result, studies that have focused on informational conformity have found that it is taken into account, by adults and children, in a broadly rational manner : people tend to be more influenced by larger groups, by stronger majorities, and when they are less sure of their own opinions (Bernard, Harris, Terrier, & Clément, 2015; R. Bond, 2005; Campbell & Fairey, 1989; Gerard, Wilhelmy, & Conolley, 1968; McElreath et al., 2005; Morgan, Laland, & Harris, 2015; Morgan, Rendell, Ehn, Hoppitt, & Laland, 2012)."

Mercier 2017

107 *"The studies mentioned above also show that the rejection of weak arguments does not stem from a blanket rejection of all arguments that challenge one's prior beliefs. Strong arguments are rated positively, and they influence participants, even when they run against prior preferences or beliefs (e.g., Petty & Cacioppo, 1979; Trouche et al., 2014; Trouche, Shao, & Mercier, 2017)...*

The experimental psychology results reviewed above demonstrate that people are endowed with mechanisms of epistemic vigilance that work, in the laboratory at least, reasonably well. They evaluate messages based on their content, on various attributes of their source, and on the arguments provided to support them: they are, broadly, ecologically rational."

Mercier 2017

108 *"...most instances of gullibility are the outcome of content-based, rather than source-based, processes, and that they only indirectly bear on the working of epistemic vigilance."*

Mercier 2017

109 *"Bloodletting is a salient example. One of the most common therapies for significant portions of Western history, inefficient at best and lethal at worse, it seems to have owed its success to the authority granted the writings of Galen and other prestigious physicians (Arika, 2007; Wootton, 2006). This thus seems to be a blatant example of misplaced prestige bias (Henrich & Gil-White, 2001). However, ethnographic data reveal that bloodletting is a common practice worldwide, occurring in many unrelated cultures, on all continents, including in many cultures which had not been in contact with Westerners (Miton, Claidière, & Mercier, 2015). These ethnographic data, as well as some experimental evidence, suggests that bloodletting owes its cultural success to its intuitiveness: in the absence of relevant medical knowledge, people find bloodletting to be an intuitive cure (Miton et al., 2015). If this explanation is correct, trust would flow in the other direction: instead of bloodletting being practiced because it is defended by prestigious physicians, it is because some physicians practiced and defended bloodletting that they became prestigious. This explanation could easily be extended to the most common forms of therapies in premodern cultures, which all aim at removing some supposedly bad element from the body (laxatives, emetics, sudation, see, Coury, 1967). Similarly, people would not refuse to vaccinate their children because they follow Jenny McCarthy or other prominent antivaxxers. Instead these figures would become popular because they attack a very counterintuitive therapy (Miton & Mercier, 2015). This phenomenon*

would thus be similar to that of political or religious leaders who are mostly deemed charismatic and prestigious because they endorse popular positions. In neither case would people be gullibly following prestigious leaders, instead they would simply be heeding messages they find appealing, and then conferring some prestige on those who defend them. The spread of misguided beliefs would thus mostly rest on content- based rather than source-based processes."

Mercier 2017

[110] *"Richard Dawkins (1976) has proposed a biological metaphor that also assumes that ideas compete but that does not assume they compete solely based on truth. Dawkins pictured culture as being composed of many individual units (the cultural analogue of genes) that undergo variation, selection, and retention. As a label for this cultural gene equivalent, he proposed the term meme. Dawkins's memes do not compete solely on truth—consider annoying commercial jingles or a chain letter that threatens doom if it is not reproduced and spread (Dawkins, 1976). In this article we follow Dawkins in explaining how ideas propagate using a variation-selection-retention approach, so to acknowledge our theoretical approach and unit of analysis, we often use the term meme for ideas that propagate in the social environment."*

Heath et al. 2012

[111] *"Cultural evolution is a vibrant, interdisciplinary, and increasingly productive scientific framework that aims to provide a naturalistic and quantitative explanation of culture, in both human and non-human species (Mesoudi 2011; Richerson and Christiansen 2013). 'Culture' is commonly defined as the body of information that is transmitted from individual to individual via social learning (rather than genetically), and colloquially includes such phenomena as attitudes, beliefs, knowledge, skills, customs and institutions. Inspired by pre-existing population genetics tools, the mathematical models of cultural dynamics developed by Cavalli-Sforza and Feldman (1981) and Boyd and Richerson (1985) first established that cultural change can be modelled as an evolutionary process yet one that is not slavishly identical in its details to genetic evolution. Today, while maintaining a solid modelling core (e.g. Kendal et al. 2009; Rendell et al. 2010; Aoki et al. 2011; Lewis and Laland 2012; Aoki et al. 2012; Kempe et al. 2014), a wide range of methodologies are used in the field of cultural evolution, including phylogenetic analysis (e.g. Gray and Jordan 2000; Tehrani and Collard 2002; Lycett 2009; Currie et al. 2010; Tehrani 2013; O'Brien et al. 2014), laboratory experiments (e.g. Mesoudi et al. 2006; Caldwell and Millen 2008; Mesoudi and O'Brien 2008; Kirby et al. 2008; Morgan et al. 2012; Derex et al. 2013; Muthukrishna et al. 2014; Tamariz et al. 2014), ethnographic field studies (e.g. Guglielmino et al. 1995; Henrich and Henrich 2010; Mathew and Boyd 2011; Hewlett et al. 2011; Demps et al. 2012; Kline et al. 2013), quantitative analysis of pre-historical, historical, and contemporary datasets (e.g. Shennan and Wilkinson 2001; Henrich 2001; Kline and Boyd 2010; Collard et al. 2011; Turchin et al. 2013; Acerbi and Bentley 2014; Beheim et al. 2014), and comparative studies of culture across species (Whiten et al. 1999; Laland et al. 2011; Dean et al. 2012). Although varied in methodology and topic, these studies are united by the notion that culture evolves according to broadly Darwinian principles. In parallel with this approach, a group of cognitive anthropologists have advanced a similar project aiming towards naturalistic explanations of culture, mainly focusing on the role that cognitive factors play in the transmission and transformation of cultural representations (Sperber 1996; Atran 1998; Boyer 2001; Sperber and Hirschfeld 2004). This approach has generated findings using laboratory experiments (e.g. Boyer and Ramble 2001; Barrett and Nyhof 2001; Norenzayan et al. 2006; Fessler et al. 2014) and analyses of historical (e.g. Nichols 2002; Norenzayan et al. 2006; Morin 2013) and cross-cultural (e.g. Atran 1998) datasets. The two approaches initially developed separately and, despite a series of attempts at seeking common ground (Henrich and Boyd 2002; Claidière and Sperber 2007; Sperber and Claidière 2008; Henrich et al. 2008), there is remaining disagreement (see e.g. Claidière et al. 2014). This disagreement rests, at a general level, in a different view of cultural transmission. For the standard cultural evolution approach, typified by Boyd, Richerson, Henrich and others, it is common to think of cultural evolution as a process of selection between different variants (e.g. beliefs, ideas or artefacts) or models (referring to people from whom one can copy). When deciding a name for a newborn, for example, one chooses from a pool of variants—the existing names in the population—and the individual-level processes of selection determine the success, at the population-level, of the variants. Cultural transmission has relatively high fidelity, and selection between faithfully transmitted variants plays an important role in determining cultural trajectories. Sperber, Claidière, Atran, Boyer and colleagues, instead, argue that in the vast majority of cases cultural traits are neither properly copied or selected, but reconstructed each time an instance of transmission happens. The permanence of some cultural traits occurs not due to high fidelity cultural transmission but instead due to the existence of stable "cultural attractors" (Sperber 1996). For example, in an oral transmission of a*

story, say Cinderella, it is highly unlikely the story will be repeated verbatim at each passage. Still, some defining features, say the pumpkin coach or the wicked stepmother, perhaps because they are particularly memorable, will act as attractors, and will be repeated ('reconstructed') each time by different narrators. Cultural transmission here has relatively low fidelity, and non-random distortions and reconstructions play an important role in maintaining cultural diversity and stability.

This general divergence has a series of consequences, ranging from what are considered the most important or interesting factors to take into account when explaining the permanence and diffusion of cultural traits (cognitive transformation of representations for Sperber and colleagues, interaction of simple decision-making biases with populational dynamics for the standard cultural evolution approach) to how far the analogy between cultural and biological evolution should be pushed (less for the former than for the latter approach)."

Acerbi & Mesoudi 2015

[112] *"Overall, the analysis presented here suggests that one of the factors that could explain the success of online misinformation is that it appeals to general cognitive preferences. Consistent with previous research, 'suspect' articles were found heavily leaning towards negative content. The various cognitive factors coded were present to a different degree. Descriptions of threats were prominent, with almost 30% of the articles containing them. Elements eliciting disgust and sexual details were also present, but they were generally co-occurring with threat-related information (the single most successful 'fake news' in Facebook in 2017 is a good example of this combination: Babysitter transported to hospital after inserting a baby in her vagina, BuzzFeed, 2017). … Articles with minimally counterintuitive elements were less common than articles with threat-, sex-, and disgust-related information. In addition, violations of intuitions that could be considered 'supernatural' in the common sense of the term were even less, making for around 5% of the articles (the other articles consisted in violations of essentialist intuitions, see below). This is partly surprising, giving the importance given in the cultural evolution literature to MCI elements. … Social information and presence of celebrities were the elements quantitatively most important. …*

This analysis suggests a few general considerations on the spread of online misinformation. First, articles concerning political misinformation, while abundant, were still technically a minority in the sample considered (40% of the articles). Different sampling methods could, of course, give different results, but this figure is consistent with the idea that online misinformation is not necessarily political misinformation. While there may be good reasons to focus the attention to the possible risks that the spread of political 'fake news' online entails, it may also be conceivable that the danger of misinformation online has been overstated by previous research, by artificially limiting the breadth of the phenomenon on explicitly malicious political articles (similar conclusions on the overestimation of the effect of political misinformation are reached, for example, in Allcott and Gentzkow, 2017 and Guess et al. 2018)…

Second, is there any specificity of the spreading of online misinformation? Various reasons have been proposed to explain why misinformation should thrive online (as opposed to offline), including the fact that everybody can quickly and cheaply spread information, that digital media make easier to find other individuals confirming incorrect information, that online interactions can preserve anonymity (Allcott and Gentzkow, 2017) and that search engines, and especially social media algorithms, are optimised for (shallow) engagement, giving disproportionate weight to 'like' and previous traffic (Chakraborty et al. 2016). This analysis points to the fact that, however, the same features that make urban legends, fiction, and in fact any narrative, culturally attractive also operate for online misinformation. While this does not exclude that specific mechanisms favour online spreading of misinformation, it suggests that to better understand them, some knowledge of why some narratives are attractive and others are not can be useful."

Acerbi 2019a

[113] *"All this show the paradox of junk culture in even starker relief. Psychologists have gathered large amounts of evidence for a series of cognitive systems to acquiring useful, that is fitness-relevant information about the world, especially from conspecifics, and ensuring that information is of sufficient quality. This seems to be a straightforward consequence of cognitive evolution. In the same way as our visual system is designed to use available information from light reflectance, our inference systems should be designed to acquire reliable information, as every increment in that capacity does translate as a survival advantage. So, again, why would humans blithely fill their minds with poor-quality information, which in most cases is of no clear advantage?"*

Boyer 2018

---

114 *"This does not mean that plausibility checking or reasoning cannot fail in their own right. People could be tricked by sophistry and various other types of manipulative messages to accept messages that are in fact largely inconsistent with their beliefs (see, e.g., Maillat & Oswald, 2009; Stanley, 2015)."*

Mercier 2017

115 *"Rumors are about mostly negative events and their sinister explanation. They describe people intent on harming us or who have already done so. They describe situations that will lead to disaster if no action is taken. The government is involved in terrorist attacks against the population, medical authorities conspire to spread mental illness in children, ethnic others are trying to invade us, and so forth. in other words, rumors describe potential danger and the many ways in which we could all be threatened. "… Human minds comprize specialize systems for threat detection. it is an evolutionary imperative for all complex organisms to detect potential dangers in their environment and engage in adequate precautionary behaviors… Threat-response systems, in humans as in other animals, face the problem that there is an important asymmetry between danger cues and safety cues. The former are actual properties of the environment… There is however no clear signal of non-danger… In humans, whose behavior is strongly affected by information from conspecifics this asymmetry of threat and safety has one important consequence, that precautionary advice is rarely put to the test. Indeed, it is one of the great advantages of cultural transmission that it spares individuals from systematically testing their environments to identify sources of danger… This would suggest that threat-related information is often considered credible, at least provisionally, as a precautionary measure. The psychologist Dan Fessler tested this directly by measuring people's acceptance of statements phrased in either negative, threat-related terms (such as "10% of heart attack patients die within 10 years") or positive terms ("90% of heart-attack patients survive for more than 10 years"). Ever though the statements are strictly equivalent, people place more confidence in the negatively framed ones. Similarly, people find the authors of descriptive texts, for example, about a computer program or a hiking trip; more competent and knowledgeable if the texts include threat-related information… So we should expect that people are particularly eager to acquire threat-related information. Naturally not all such information could give rise to rumors that people take more seriously than mere urban legends, otherwise cultural information would consist in nothing but precautionary advice. But several factors limit the spread of rumors about potential threats. First, all else being equal, plausible warnings have an advantage over descriptions of highly unlikely situations. … It is generally easier to convince our neighbors that the grocer sells rotten meat than that he occasionally turns into a reptile. Note that, as a matter of course, what is or is not plausible depends on the listener's own metrics… Second, the niche for non-tested (and generally invalid) precautionary information requires that the cost of precautions be relatively moderate. To take an extreme case, it is relatively easy to convince people not to walk around a cow seven times at dawn… Third, the potential cost of noncompliance, what would happen if we failed to take precautions, should be described as serious enough that the listener's threat-detection systems are activated… So it would seem that threat detection is one of the domains in which we may have to turn down our epistemic vigilance mechanisms and take as a guide to behavior precautionary information, especially of it is not too costly to follow, and if the averted danger is both serious and uncertain."*

Boyer 2018

116 *"Deception may be adaptive, if you can exploit others, but then it becomes adaptive for others to develop the symmetrical weapon, the ability to see through deception. There is an equilibrium when capacities for deception and detection are roughly equivalent. But that equilibrium is unstable. Any organism that is slightly better than others at deception will gain an advantage, so that it will transmit its deceptive skills to its offsprings, until these skills become the population average. But then an increase in detection skills become adaptive, and in a similar way will gradually become the average. This kind of arms race between deception and detection is common in nature. In the case of human communication, the arms race consists in competition between the capacity to make one's utterances persuasive, on the one hand, and the ability to protect one's own beliefs from deception, on the other."*

Boyer 2018

117 *"There is also a motivation to transmit, without which many people would cultivate their own poor-value information but there would be no rumors, no junk culture.*

In many situations the transmission of low-value information is associated with strong emotions. People consider information about viruses and vaccinations and government conspiracies as terribly important. When *they transmit information about such topics, people are not just eager to convey but also eager to convince. They do pay attention to their audiences' reactions, and they consider skepticism highly offensive. Doubt is attributed to all sorts of wicked motives… Why are the beliefs so intensely moralized? One obvious answer is that the moral value of broadcasting the information, and of accepting it, is a straightforward consequence of the information conveyed… A crucial part of our evolved psychology consists in capacities and motivations for efficient coalition management. So, when humans convey information that may persuade others to engage in specific actions, we should try to understand this in terms of coalition recruitment. That is to say, we should expect that an important part of the motivation here is indeed to persuade others to join in some collective action. … Roughly speaking, stating that someone's behavior is morally repugnant creates consensus more easily than claiming that the behavior results from incompetence. The latter could invite discussions of evidence and performance, more likely to dilute consensus than to strengthen it."*

Boyer 2018

[118] *"To sum up, even from the earliest stages of cognitive development human minds seem to be designed to acquire useful knowledge about their environment. I must insist on the word 'useful'. We should not assume that human minds are designed to acquire true information about their natural and social environments. That is an important difference. Just because something is a fact does not mean that humans are equipped to find out about it. Conversely many of our intuitive expectations lead us to false beliefs. … So it is important to remember that the human mind is not always philosophically correct or scientifically accurate. The assumptions it contains may not be true, but they are useful. Usefulness, then, refers to selective pressure."*

Boyer 2018

[119] *"The evolution of the human cultural capacity – that is, for intergenerationally stable, high fidelity, social transmission – created a new selective environment in which mutations improving the reproductive benefits of such transmission were favored. Our ancestral psychology evolved within physical and phylogenetic constraints) into an increasingly well-organized and specialized battery of biases jointly designed to extract reproductive benefit from the flow of socially transmitted information. Prestige processes emerge from this evolved social learning psychology. Cultural transmission is adaptive because it saves learners the costs of individual learning. Once some cultural transmission capacities exist, natural selection favors improved learning efficiencies, such as abilities to identify and preferentially copy models who are likely to possess better-than-average information. Moreover, selection will favor behaviors in the learner that lead to better learning environments, e.g., gaining greater frequency and intimacy of interaction with the model, plus his/her cooperation. Copiers thus evolve to provide all sorts of benefits i.e., 'deference') to targeted models in order to induce preferred models to grant greater access and cooperation. Such preferred models may be said to have prestige with respect to their 'clients' the copiers). The above implies that the most skilled/knowledgeable models will, on-average, end up with the biggest and most lavish clienteles, so the size and lavishness of a given model's clientele the prestige) provides a convenient and reliable proxy for that person's information quality. Thus, selection favors clients who initially pick their models on the basis of the current deference distribution, refining their assessments of relative model worth as information becomes available through both social and individual learning. This strategy confers a potentially dramatic adaptive savings in the start-up costs of rank-biased social learning. Finally, because high-quality information 'expertise,' 'performative skills,' 'wisdom,' 'knowledge') brings fitness-enhancing deferential clients, models have an extra incentive to outexcel each other."*

Heinrich & Gil-White 2001

[120] *"Participants watched an 'attentional cuing' clip, where two models received unequal bystander attention. In this cuing scene two bystanders stood between the models, attending to only one of them — the 'prestigious model.' This prestige cuing was followed by four 10-seconds (s) 'test' clips, where those two models demonstrated different behaviors, preferences and labels. In all test clips solitary models demonstrated their preference towards an object; then participants' own preferences toward those same stimuli were recorded. The order in which models appeared and the identity of the prestigious model were counter-balanced across participants...*

Our findings provide support for the existence of a domain-sensitive prestige bias in children's learning: children's learning from cultural models was biased by the mere preferential attention of bystanders, particularly on activities similar to those the model had been engaging in when she received bystander attention. These strong effects from a minimal manipulation suggest that prestige bias may be a potent pressure on *cultural evolution. As predicted (Henrich & Gil-White, 2001), we witnessed biased learning in different domains, including potentially costly dietary preferences…*

With regard to cross-domain effects, Henrich and Gil-White (2001) predicted that prestigious individuals are 'influential, even beyond their domain of expertise.' Recent developmental research (Fawcett & Markson 2010) has indicated that 2-year-olds who know a model shares their preferences in one domain (food or television shows) will only imitate that model's preferences in that same domain, not the other. This suggests that mere 'similarity cues' may not be strongly influential beyond their domain of expertise. In our work with 'prestige cues', we witnessed an interesting domain-based asymmetry. Our subjects' food and drink preferences trended toward prestige-bias when they saw artifact-use cues (in study 1, the combined food and drink measures registered a significant effect after a prestige cue, but each measure independently did not); however, their artifact-use preferences trended away from prestige-bias after seeing food cues. This raises the interesting possibility that children's inferences about model quality exploit an asymmetric map of the relationships between learning domains."

Chudek et al. 2011

[121] *"The evidence reviewed in this article provides mixed support for the use of prestige-biased social learning in both human adults and children. However, few studies have examined this and further research is needed to clarify which factors lead to variation in the use of prestige-biased social learning. The difficulty of the task, the relevance of the domain for the individuals and the benefits and costs associated with the task seem to be important factors influencing the use of prestige-biased social learning (see prediction (viii) in Table 2). In general, easy tasks, tasks that are not relevant for participants and tasks that do not provide incentives to perform well or avoid costs (e.g., monetary rewards or costs) seem not to stimulate the use of prestige-biased social learning (Acerbi and Tehrani, 2018). Other factors taken into account in the literature, such as experience and age (Little et al., 2015) seem to be important when they affect task difficulty, the relevance of the domain for the participants and potential gains or costs of the task for the participants. For instance, expertise leads to a greater use of prestige-biased social learning when the task is more relevant for the experts but the task is still difficult for them (Verpooten and Dewitte, 2017). Similarly, younger individuals use more prestige-biased social learning than older individuals when the task is more relevant for them (Little et al., 2015). Moreover, when there is little variation in knowledge/skill in a group, it is more adaptive to learn from low access cost models than from costly prestigious models (Henrich and Henrich, 2010; see prediction (ix) in Table 2). Another factor that influences the use of prestige-biased social learning is the availability of alternative social learning biases, e.g., success or content biases. When success information is provided, this information should be preferentially used over prestige information (prediction (ix) in Table 1). However, this was not found in the sole experiment comparing prestige with success bias (Atkisson et al., 2012), although this is a single study. Both direct and conceptual replications are needed to gain confidence in this result. Content bias was stronger than prestige bias in another study (Acerbi and Tehrani, 2018), but this might depend on the domain and the factors mentioned above (i.e., task difficulty, relevance for the individual, and benefits and costs associated with the task). Variation in some of these factors (e.g., the relevance for the participants) might lead some participants (e.g., non-experts) to make use of content biases, while other participants (e.g., experts) to employ prestige-biased social learning (Verpooten and Dewitte, 2017). It is also possible that prestige biased social learning has different effects on different measures of influence, e.g., recall, likability, behavioural influence, willingness to transmit and receive information. To the extent these measures of influence affect task difficulty, relevance for the participants or the benefits/costs associated with tasks, it seems plausible that the different measures would be a source of variation in the use of prestige and other social learning biases. For instance, although one recent study found that anti-vaccination messages are not better transmitted per se, exploratory analyses showed that when anti-vaccination messages are provided by doctors (i.e., a prestigious source within a relevant domain) these types of messages are especially powerful in influencing people's vaccination-related decisions (Jiménez et al., 2018). Similarly, although people might be able to appreciate the content of certain pieces of information (e.g., quotes, news, artworks, etc.) independent of the prestige of the source of the information, they might be more influenced by prestige cues when they want to achieve influence over other people's behaviour (e.g., by quoting a prestigious source of information), get personal or social benefits (e.g., choosing artworks to be displayed in their own town) or they have to decide whether to learn more about a topic or transmit the information about the topic to other people. Therefore, research on prestige-biased social learning might benefit from comparing the influence of prestige cues on different types of outcomes…*

In conclusion, H&GW's theory of the evolution of prestige has generated a great deal of research and this research has stimulated new research questions and predictions. Although the evidence reviewed here suggest that prestige-bias social learning is employed in at least some contexts, further research will need to determine the precise circumstances in which people use prestige cues to learn socially, and when the use of these cues is adaptive."

Jimenez & Messoudi 2019

[122] *"In our studies we presented respondents with questions that had real consequences for them as decision makers, since they received a bonus for making accurate judgments. The respondents were given advice and the principal measure was the weight placed on that advice in their final decisions. The studies, which were conducted on a computer due to their interactive nature, shared the following general procedure. In the first phase, respondents were presented with questions and asked to state their estimates. In the second phase, they were presented with the same questions along with estimates made by various advisors (other students). The respondents were then asked to provide their estimates once again. They were free to use the advice as they wished… A coherent picture emerges from the advice weighting policies observed across the studies. First, the results of Study 1 show egocentric discounting of advice. Second, advice discounting was not indiscriminate; individuals had a veridical view of their knowledge, so that the less knowledgeable ones placed greater weight on the advice (Studies 1–3). Third, the weight of advice declined with the distance between the advice and their initial opinions (Studies 2–3); this distance effect was exhibited in the high-knowledge condition and to a lesser extent in the low-knowledge condition as well. … The asymmetric weighting of one's own and other opinions is attributed to the fundamental asymmetry in access to the underlying justifications for each opinion. Decision makers can assess what they know and the strength of their own opinions, but are far less able to assess what an advisor knows and the reasons underlying her/his opinions. Naturally, one's confidence about a given opinion (or hypothesis) is related to the amount of evidence that one could readily recruit to support it. Other things being equal, decision makers are likely to feel more confident about their own opinion than about the other opinion, hence their own estimate would receive greater weight than the advice. Earlier findings suggest that respondents weight each opinion according to the expertise ascribed to its source (Birnbaum & Stegner, 1979; Birnbaum & Mellers, 1983). The self/ other asymmetry presumably enhances the expertise ascribed to the self. This line of reasoning about information asymmetry is also reminiscent of the principal- agent problem in organizations (Eisenhardt, 1989).*

The explanation of the self/other effect in terms of differential information access seems preferable to alternative explanations that posit either a self-serving bias (e.g., an optimistic bias) or commitment to one's past decisions as the root of discounting others' views. To be sure, self-serving biases pervade interpersonal comparisons, in that, for example, people believe that they have lower chances of experiencing negative life events, such as car accidents and strokes, than others do or that they rank higher than others on various abilities and attributes, such as driving ability and social skills (e.g., Brown, 1986). But a bias of this sort does not readily explain respondents' weighting policies for advice, especially the sensitivity of those policies to the respondents' own knowledge (Studies 1–3) and their sensitivity to the quality of the advice (Yaniv & Klein- berger, 2000).

Commitment to one's past decisions is a powerful motive in decision-making, yet it cannot readily explain the findings either. The antecedents of commitment—high costs for being inconsistent, the need to justify decisions to others, having to admit past mistakes, and having to save face with respect to ego-involving issues—were largely absent in the present studies. Our respondents made their judgments in a private setting (by entering responses into a computer file), received incentives for accuracy, and were not asked to justify their estimates.

A cognitive explanation based on informational asymmetry and the assessment of available evidence is more parsimonious and hence superior to those based on a self-serving bias or commitment because it can readily account for the finding that respondents' weights on advice are sensitive to the quality of the advice (Yaniv & Kleinberger, 2000) as well as their own knowledge (e.g., Study 1), without making unnecessary assumptions."

Yaniv 2004

[123] *"The literature on social cognition emphasizes infants' emerging understanding of other people as repositories of knowledge who can adopt varying cognitive and emotional perspectives toward objects and events (e.g., Baldwin & Moses, 1996; Moses, Baldwin, Rosicky, & Tidball, 2001; Tamis-LeMonda & Adolph, 2005). With the realization that caregivers can provide useful social information, infants increasingly turn to them for guidance when they are unsure about how to respond. Infants' capacity to benefit from others' social*

advice enormously expands opportunities for learning because infants need not rely solely on learning through self-discovery; now infants can seek others' advice in ambiguous situations and respond to unsolicited advice about how to act (Moses et al., 2001)."

Tamis-LaMonda et al. 2008

[124] *"Cultural evolution theory posits that a major factor in human ecological success is our high-fidelity and selective social learning, which permits the accumulation of adaptive knowledge and skills over successive generations. One way to acquire adaptive social information is by preferentially copying competent individuals within a valuable domain (success bias)."*

Jimenez & Messoudi 2019

[125] *"Cognition refers to all the information processing carried out by the brain, particularly those involved in learning at school. Metacognition refers to the cognitive processes that control and evaluate cognition itself. In other words, the term "metacognition" refers to the set of processes by which each of us regulates our attention, chooses to be informed, plans, solves a problem, identifies errors and corrects them. At school, this set of abilities plays a central role. Good "regulation" leads the student to engage in learning with confidence and enthusiasm. Poor regulation" of metacognition results in dislike of learning, avoidance of school, dropping out, and what is known as the "spiral of failure".* »

Proust 2019

[126] "*Metacognition is a broad term, and often interpreted differently by different researchers. As a first step, it is crucial to separate the empirical definition of metacognition from its epistemological status as a meta-level representation of an object-level cognition. Empirically, metacognition is often operationalized as 'behaviour about behaviour' rather than 'cognition about cognition' (see table 1 in Fleming & Dolan). Here, we define second-order behaviours as decisions contingent on other behavioural outputs (that either have occurred or will occur). Consider a visual detection task. Following a first-order response as to whether the stimulus is present or absent, a confidence judgement in one's response being correct is second-order with respect to the previous decision. This does not necessarily entail that the second-order judgement requires a meta-level representation of the object-level decision; it could instead be accomplished via object-level representations, for example, by basing confidence on information about the stimulus. Alternatively, the confidence judgement could be based on a meta-representation of the decision and subsequent response. This creates an initial division of the theoretical landscape, with two orthogonal dimensions—those of level of representation and order of behaviour."*

Fleming Dolan & Frith 2012

[127] *"Whether it is a physical or cognitive action, the brain does not let us engage in an activity without anticipating our chances of success. Experimental work shows that the primary source of information used by metacognitive self-assessment is the predictions of success or failure provided by the feelings that are produced during the activity. They are very different depending on the cognitive activity and the segment of the activity they concern (e.g. its future feasibility or final correction).* »

Proust 2019

[128] *"How does the brain select its predictive cues, i.e. the success criteria for a given activity? It selects them by reinforcement. Thanks to observations of neural activity in animals, we now know some of the predictive cues used by the brain. For example, how quickly the brain begins to activate the processing of task-related information is a predictive indicator. Other indicators are related to the way in which coordination between neuronal assemblies takes place, in particular how quickly they converge towards a single decision. In both cases, the cues predicting success - for example, the speed of neuronal activation or the speed of convergence towards a single decision - are compared with the cues observed in the present context. The set of indices together form what are called "predictive heuristics" (see Figure 4). We don't know which heuristics our brain has used, but we experience the resulting feelings. It is these feelings that allow us, at any given moment, to know what we know, what we understand, what we want to learn, and what we have succeeded or failed to do.* »

Proust 2019

[129] "*… Hampton (Hampton, 2001) devised a prospective memory confidence task. Trained monkeys performed a delayed-match-to-sample task. In this task, an image, referred to as the target, appeared at the beginning of a*

trial. At the end of the trial, after a delay, animals were required to select the target that reappeared with another series of images (distractors). To evaluate if the monkeys remembered the target, after two thirds of the delay, the monkeys received an option to accept the test or decline it. If they accepted it, and they made a correct match, they received a large reward. However, if they made a mistake by choosing a distractor as a match, then they received no reward. If the monkeys declined the test, they received a small reward, regardless of whether they chose the target or a distractor as the match. The investigators reasoned that if the monkeys believed they would perform well, they would accept the test, choose the correct target and receive a large reward. However, if they were uncertain, they would decline the test, and opt for the small but certain reward. Therefore, in this task, the monkeys made a prospective judgment about how they were likely to perform on the test. An additional strength of the task design was that four stimuli were used as possible targets and were selected as targets randomly each day. Because the stimulus set changed across sessions, the monkeys could not associate one particular stimulus with the likelihood of correct responding. Rather, they had to rely on their memory of the sample stimulus to decide whether to accept or decline the next step of the task. In addition, the monkeys could not use cues such as their own reaction time to estimate the likelihood of a correct response, because they had to decide to take the test or not before the match choice was required. As predicted, monkeys opted out when they did not remember, consistent with the hypothesis that they were less confident on those trials. Indeed, they performed better in this task then in a similar forced task, when they did not have the option to decline the test.

In a similar spirit, Son and Kornell (2005) trained rhesus macaque monkeys to distinguish the length of two lines. After the monkeys made their decision, consisting of choosing the longest line, they were required to rate their confidence in their decision by making a bet, that is, a retrospective task. Two betting options were represented by two choice targets. If the monkeys chose the low bet target, they received a small reward, regardless of whether their previous response on the discrimination task was right or wrong. If they chose the high bet target, they received a large reward for correct responses and no reward for incorrect responses. Monkeys generally chose low rewards more frequently in difficult discrimination trials indicating that they knew when they did not know. The same monkeys engaged in the same betting strategy during a dot-density discrimination task, showing that they could generalize their reports of confidence to different tasks. Similar approaches have been used to study confidence in smaller mammals such as rodents. Foothe and Crystal (2007) trained rats to discriminate the duration of sounds. In each trial, the rats were able to choose if they wanted to take a test or not. Similar to the monkeys, rats chose to avoid the test when the stimulus was ambiguous."

Grimaldi, Lau & Basso 2015

[130] "*Often, researchers restrict the definition of metacognition to the kind of second-order behaviour available for subjective report [6]. In many of these cases, it seems intuitive that the report is capturing an aspect of cognition that is secondary to the cognitive process itself. Take the case of blindsight [9]: in some patients with lesions to primary visual cortex, visuomotor performance when responding to targets in the 'blind' field may be well above chance, yet the patient reports not seeing anything. This is a case where first-order (visuomotor) performance is high, but awareness is absent. Yet, the reliance on subjective reports to index metacognition precludes the ascription of metacognition to non-human animals and non-verbal infants, and may prematurely equate metacognition with consciousness (see §6). In contrast, non-verbal behavioural measures do not suffer from these drawbacks. Smith et al. review a large body of work in non-human animals using the 'uncertain-option' paradigm… whether the second-order behaviours often used to index metacognition can be explicable in non-metarepresentational terms remains an empirical question. We might find that an object- level account is sufficient to explain second-order behaviour in some circumstances, but not others. On the other hand, evidence from human neuropsychology that first- and second-order behavioural performances are dissociable suggests that at least some degree of separate representation will be required to account for second-order behaviours."*

Fleming, Dolan & Frith 2012

[131] *"Metacognitive confidence can be formalized as a probability judgment directed toward one's own actions—the probability of a previous judgment being correct. There is a rich literature on the correspondence between subjective judgments of probability and the reality to which those judgments correspond. For example, a weather forecaster may make several predictions of the chance of rain throughout the year; if the average prediction (e.g., 60%) ends up matching the frequency of rainy days in the long run we can say that the forecaster is well calibrated. In this framework metacognition has a normative interpretation as the accuracy of a probability judgment about one's own performance."*

Fleming & Lau 2014

132 "*In studies of human perceptual decision-making, confidence is often measured with retrospective judgment. Subjects give a confidence rating right after a report about a perceptual experience and therefore must base their confidence judgment on the memory of their initial response. For example, a subject might first perform some perceptual task such as reporting their perception of an ambiguous object (do you see a vase or a face?). Then the subject would immediately declare how confident s/he felt about that decision.*

Similar to measures of confidence using open-ended ratings, several scales have been developed to measure confidence more quantitatively. The most commonly used is confidence rating. In this scale, the subject is asked to report confidence on a continuous scale ranging from 0% or complete uncertainty to 100% or complete certainty. Alternatively, it can be assessed with discrete fixed levels, or a simple binary choice (confident/not confident, Cheesman and Merikle, 1986; Dienes and Perner, 1999). However the use of ratings has been criticized because some subjects may find it not intuitive or they may be poorly motivated to accurately report their confidence (Persaud et al., 2007). To overcome these limitations, post-decision wagering has been introduced, in which subjects bet money or tokens on their own decisions (Persaud et al., 2007; Ruffman et al., 2001). In this context, subjects should ideally bet low when they are not confident and bet high when they are confident, in order to maximize gain. This task is more engaging and more intuitive for most participants. However, it has been noted that wagering can be influenced by individual propensity to risk (Fleming and Dolan, 2010) and that subjects tend to use only the ends of the scale, probably in order to maximize gains (Sandberg et al., 2010), thus suffering from low sensitivity for intermediate ranges. In an attempt to develop a scale that has both the sensitivity of confidence ratings and the intuitiveness of post decision wagering, the feeling of warmth scale has been developed (Metcalfe, 1986; Wierzchon et al., 2012). In this scale subjects report their confidence as a temperature, ranging from cold (not confident) to hot (very confident), with intermediate options (e.g. chilly or warm). The perceptual awareness scale (Ramsøy and Overgaard, 2004) and the Sergent-Dehaene scale (Sergent and Dehaene, 2004) are also commonly used and were developed to judge the degree of visibility in visual tasks, ranging from no visibility at all to clear perception, with discrete intermediate levels (perceptual awareness scale), or a continuous spectrum (Sergent- Dehaene scale). When applied to confidence, however, these two scales end up being very similar to confidence rating. An extensive discussion of the properties and sensitivities of the different scales is beyond the scope of this review. For a rigorous comparison see, Sandberg et al., (2010) and Wierzchon et al., (2012)."

Grimaldi, Lau & Basso. 2015

133 "*Hampton (Hampton, 2001) devised a prospective memory confidence task. Trained monkeys performed a delayed-match-to-sample task. In this task, an image, referred to as the target, appeared at the beginning of a trial. At the end of the trial, after a delay, animals were required to select the target that reappeared with another series of images (distractors). To evaluate if the monkeys remembered the target, after two thirds of the delay, the monkeys received an option to accept the test or decline it. If they accepted it, and they made a correct match, they received a large reward. However, if they made a mistake by choosing a distractor as a match, then they received no reward. If the monkeys declined the test, they received a small reward, regardless of whether they chose the target or a distractor as the match. The investigators reasoned that if the monkeys believed they would perform well, they would accept the test, choose the correct target and receive a large reward. However, if they were uncertain, they would decline the test, and opt for the small but certain reward. Therefore, in this task, the monkeys made a prospective judgment about how they were likely to perform on the test. An additional strength of the task design was that four stimuli were used as possible targets and were selected as targets randomly each day. Because the stimulus set changed across sessions, the monkeys could not associate one particular stimulus with the likelihood of correct responding. Rather, they had to rely on their memory of the sample stimulus to decide whether to accept or decline the next step of the task. In addition, the monkeys could not use cues such as their own reaction time to estimate the likelihood of a correct response, because they had to decide to take the test or not before the match choice was required. As predicted, monkeys opted out when they did not remember, consistent with the hypothesis that they were less confident on those trials. Indeed, they performed better in this task then in a similar forced task, when they did not have the option to decline the test.*

In a similar spirit, Son and Kornell (2005) trained rhesus macaque monkeys to distinguish the length of two lines. After the monkeys made their decision, consisting of choosing the longest line, they were required to rate their confidence in their decision by making a bet, that is, a retrospective task. Two betting options were represented by two choice targets. If the monkeys chose the low bet target, they received a small reward, regardless of whether their previous response on the discrimination task was right or wrong. If they chose the high bet target, they received a large reward for correct responses and no reward for incorrect responses.

Monkeys generally chose low rewards more frequently in difficult discrimination trials indicating that they knew when they did not know. The same monkeys engaged in the same betting strategy during a dot-density discrimination task, showing that they could generalize their reports of confidence to different tasks. Similar approaches have been used to study confidence in smaller mammals such as rodents. Foothe and Crystal (2007) trained rats to discriminate the duration of sounds. In each trial, the rats were able to choose if they wanted to take a test or not. Similar to the monkeys, rats chose to avoid the test when the stimulus was ambiguous.

Grimaldi, Lau & Basso 2015

[134] *"To address this issue, we combined a nonverbal memory-monitoring paradigm developed for rhesus monkeys (22) with a pointing paradigm suitable for human infants. Twenty-month-old infants (n = 80) had to remember the location of a hidden toy for a variable delay before pointing to indicate where they wanted to recover it (Fig. 1A). Task difficulty was manipulated along two orthogonal dimensions: (i) Infants had to memorize the location of the toy for a variable delay (3, 6, 9, or 12 s), and (ii) they either saw the toy being hidden at a given location (possible trials) or could not see where the toy was being hidden (impossible trials). Crucially, half of the participants were given the possibility to avoid responding by asking their caregiver for help (AFH) instead of pointing (experimental group; n = 40), whereas the other half were not given this opportunity and could only choose a location by themselves (control group; n = 40). This manipulation enabled us to test whether infants can monitor and communicate their own uncertainty. Indeed, if infants can monitor their own knowledge state, they should use the AFH option (i.e., opt-out) when they have forgotten the toy location, thereby avoiding mistakes and improving their performance …Furthermore, if infants can monitor the strength of their memory trace, they should use the AFH option more often at higher levels of uncertainty (i.e., for longer delays and impossible trials)… We then tested whether task difficulty had an impact on the probability of asking for help. Indeed, if infants were monitoring their own uncertainty about the toy location, they should have asked for help more often as the memorization delay increased. This analysis was restricted to the participants in the experimental group, who asked for help in at least one trial per condition (n = 21). An ANOVA revealed that the probability of asking for help was higher for impossible than for possible trials [Fig. 2A; $F(1,20) = 24.22$; $P < 0.001$]. Furthermore, within possible trials, the probability of producing an AFH response increased with increasing delays [Fig. 2B; $F(1,20) = 4.62$; $P < 0.05$]. Thus, infants' tendency to ask for help varied with task difficulty, suggesting that infants used the AFH option strategically to avoid responding when they felt uncertain about the toy location…*

When given the opportunity to decide whether they should respond by themselves or avoid responding by asking for help, 20-mo-olds are able to strategically adapt their behavior. That is, they selectively seek help to avoid making errors and to avoid difficult choices. In the comparative literature, these adaptive "opt-out" behaviors have been taken as evidence for metacognitive uncertainty monitoring in several species … However, some authors have argued that such behavioral patterns could also be explained by associative or reinforcement learning mechanisms … For instance, they suggest that difficult trials are simply avoided because individuals learn that the probability of obtaining a reward is lower for those trials … Whether or not this associative interpretation can be ruled out in comparative research, in which animals are extensively trained, remains a controversial issue … However, in the present study, an associative account seems unwarranted because infants only received a few trials (i.e., a maximum of two trials for each level of task difficulty), leaving little room for associative learning. Moreover, the proportion of AFH responses did not increase across time [effect of trial rank on the proportion of AFH responses: $F(1,20) = 0.22$; $P > 0.6$], ruling out an associative interpretation in terms of reinforcement learning. Another issue raised in the comparative literature concerns the fact that when the opt-out alternative is available simultaneously with another choice, some competition might take place between these options .... This might eventually lead to the opt-out option being triggered by default whenever the participant is unable to accumulate enough evidence and commit to a decision before a deadline has been reached. Under this account, infants in our study would simply ask for help by default when no memory is available to trigger an appropriate motor plan. However, if infants simply turned to their parents automatically when no response came to their mind (e.g., to seek comfort), we should observe a similar tendency in the control group. In fact, although infants in the control group were not taught that they could ask for help, and even though their caregiver remained unresponsive, we did observe a few spontaneous 'AFH-like' responses in this group [mean number of AFH responses in the control group: 0.6; in the experimental group: 1.42; $t(39) = 3$; $P < 0.005$; Fig. S3]. However, when we analyzed the frequency at which infants looked toward the parent in the control group, we found absolutely no increase with task difficulty (Fig. S3A), and excluding those trials did not impact performance (Fig. S3B). Thus, infants in the control group did not orient selectively toward their *parents when they were more likely to have forgotten the toy location. In turn, this finding confirms that infants in the experimental group did not automatically turn toward their parents when no response came to their mind. Rather, our results are consistent with the idea that infants in the experimental group learned that they could communicate with their caregiver to obtain some help whenever they felt that they were likely to make an error."*

Goupil, Romand-Monnier & Kouider 2016

[135] *"In experiment 1, 18-month-old infants (n = 29) saw an object being hidden in one of two opaque boxes and, after a delay, were asked to point to indicate where the object was concealed (see Figure 1A and the Supplemental Experimental Procedures). First-order performance on this task was assessed along a para-metric variation of difficulty (i.e., memorizing the location of the toy for a brief or longer delay). Immediately following this choice, infants were provided with the selected box. The amount of time they were willing to search within this box before giving up was used as a measure of post-decision persistence. Importantly, persistence times (PTs) were measured in the absence of any external feedback on performance, allowing us to use this measure as a proxy for confidence [7]... Consistent with our hypothesis, we observed that infants searched longer in the box following correct as compared to incorrect decisions ($t(28) = 2.1$; $p < 0.05$)."*

Goupil & Kouider 2016

[136] "*In the current neuroscience literature there is a fair amount of confusion regarding how confidence is encoded in the brain. Some data indicate that confidence may be encoded by the same circuits involved in decision-making, others that confidence is monitored by dedicated structures...*

All the above studies suggest that confidence is implemented in regions that are not commonly considered as part of the decision-making circuitry, evoking the image of a looker inside the brain. Some of these studies, like Lau and Passingham (2006), Del Cul et al., (2009), Rounis et al., (2010) Fleming et al., (2010) McCurdy et al., (2013), Komura et al., (2013) and Lak et al., (2014) show a clear dissociation between performance and confidence, others like, Hebart et al., (2014) Middlebrooks and Sommer (2012) and Kepecs et al., (2008) show correlates of confidence in regions that are not traditionally considered to be involved in decision-making. Together these results suggest that there are separate and perhaps multiple areas involved in confidence monitoring and reporting. Future studies could be aimed at elucidating how these areas work together to form the circuit involved in monitoring and reporting confidence…

Although neurological, neuropsychological, fMRI and psychophysical data described above support the idea that confidence circuitry is separate from decision-making circuitry, recent electrophysiological experiments in monkeys suggest that these circuits are shared."

Grimaldi, Lau & Basso 2015

[137] "*Imagine that you are driving your car at night. There are no street lights on the road and your car's front lights are dim. As you are trying to keep the car on the road you need to determine which direction you and the other traffic are moving. This can be achieved by processing two distinct sensory inputs: the visual flow field created on the retina by your own motion and the vestibular stimulation, which measures acceleration. If the car in front of you suddenly brakes, you have to make a quick decision, based on these sensory inputs, about whether it is better to veer left or right (we are assuming that there are no additional obstacles or cars on either side, in which case the only important question is how to avoid a collision with the car ahead). The best decision requires determining whether your current heading is to the left or right side of the braking car and then to veer in that direction. The noise in the vestibular system as well as the glare of lights and random movement of cars creates uncertainty and, given these sources of stochasticity, you, or rather your brain, cannot know for sure the precise direction of heading. As this example illustrates, to perform well, the brain needs to be effective at dealing with a daunting array of uncertainties. Some originate in the external world, such as sensory or motor variability, whereas others are internal to the brain and are associated with cognitive variables, timing or abstract states. When dealing with these uncertainties, it is useful to represent current knowledge with probability distributions and update these on the basis of the rules of probabilistic inference—namely Bayes' theorem. Notably, there is ample experimental evidence that humans and other animals can indeed estimate and employ uncertainty to perform probabilistic inference about sensory, cognitive and motor variables (see ref. 9 for a review). In fact, in the particular case of heading direction, humans and animals have been shown to perform near optimally given the uncertainty inherent to the visual and vestibular information. There is also emerging evidence about how brains implement these uncertainty-based computations in neural circuits."*

Pouget et al. 2012

138 *"The mathematical foundation of Bayesian concepts stems from the so-called Bayes' rule, named after one of its contributors, the 18th century British Reverend Thomas Bayes. Let's consider a practical example of how Bayes' rule works. A medical doctor faced with the following data D, a patient with a cough, contemplates three hypothetical diseases: a lung cancer (H_1), a cold (H_2) or gastroenteritis (H_3). The relative merit of each hypothesis can be deconstructed as follows according to Bayes' rule. Patients usually cough when afflicted by lung cancer or a cold but rarely in the case of gastroenteritis. Therefore, the likelihood of the potential cause for the cough is high under H_1 and H_2 and low under H_3. Second, a cold and gastroenteritis are much more prevalent diseases than lung cancer in the general population. The a priori likelihood of H_2 and H_3 is much higher than that of H_1. Given that only H_2 scores high both in a priori and current evidence, the most likely disease given the symptoms is a cold. Stated more generally, Bayes' rule says that our degree of belief in a hypothesis H given some current data D depends on the a priori likelihood of this hypothesis (what we know about it, independent of the current data), and the likelihood of the current data given this hypothesis. Formally, degrees of belief and likelihoods correspond to probabilities [1] and Bayes' rule reads:*

$p(H|D) = p(D|H)*p(H)/p(D)$.

Bayes' rule distinguishes between our belief a priori in the hypothesis p(H) and our belief in this hypothesis a posteriori, p(H|D), once particular data are considered to evaluate it. The notation p(D|H) is a shorthand for the probability of D given that we know H (the so-called likelihood of the data) and p(H|D) for the probability of H given that we know D."

Meyniel 2016

139 *"A theoretical explanation of how confidence is encoded by the same neurons involved in decision-making is supported by the currently popular, Bayesian views of the brain (Friston, 2012; Lau, 2008; Pouget et al., 2013). Bayes theorem is a way to quantify uncertainty and is formally stated as: P(a/b) = P(b/a) P(b)/P(a) where P(a|b) is the conditional probability of event a occurring given the occurrence of event b, also called the posterior. P(b|a) is the conditional probability of observing event b given event a. This is also known as the likelihood. P(b) is the probability of event b also referred to as the prior. P(a) is a normalization term and for explanatory purposes can be ignored...*

Thinking about the brain in Bayesian terms is somewhat intuitive. Neurons, particularly those in sensory and motor areas have tuning curves; that is they can be described as radially symmetric functions of a stimulus parameter. Neurons show maximal discharge for optimal stimuli or movements and discharge less with stimulus or movement parameters that are less than optimal (Chalupa, 2003). Tuning curves are essentially, likelihood functions. They are measures of the probability of a particular outcome given a particular discharge rate (Foldiak, 1993; Jazayeri and Movshon, 2006; Sanger, 2002, 2003). Recent theoretical work (Ma et al., 2006), supported by the previously described experimental studies (Beck et al., 2008; Kim and Basso, 2010) shows that populations of neurons representing the likelihood and the prior, can be combined linearly in much the same way as Bayes' theorem combines two probability distributions, to provide a read-out of a decision in the form a posterior distribution…

A critical feature of this kind of an approach to understanding decision-making is that confidence or uncertainty is encoded implicitly across the population response or the posterior…

This kind of view, that perceptual decisions are computed in terms of Bayesian probability distributions in the brain, is one motivation for believing that confidence does not depend on specialized circuitry. If perceptual decisions are already computed in such probabilistic terms, confidence information should already be present in the circuits for decision-making… However, one concern is whether such information in the superior colliculus or LIP for example, can be read-out by those structures themselves. Even if the information is there, it is still possible that a monitoring module reads out the width of the distribution of the posterior."

Grimaldi, Lau & Basso 2015

"*A general understanding of the notion of confidence is that it fundamentally quantifies a degree of belief, or synonymously, a degree of reliability, trustworthiness, certitude, or plausibility. This common notion coincides closely with a formal one: that of Bayesian probability. Although a probability is sometimes considered to describe the likelihood of occurrence of random events in the world, from the viewpoint of an observer, whether* *such likelihoods constitute objective facts or reflect subjective knowledge is indistinguishable. Thus, probabilities simply are degrees of belief from the Bayesian viewpoint (Jaynes, 2003). Recognizing that much remains to be unpacked, we adopt the notion of Bayesian probability as the formal definition of subjective confidence."*

Meyniel, Sigman & Mainen 2015

"*A key claim of this review is that the notion of 'uncertainty' used in research on Bayesian neural computation (Fiser et al., 2010; Ma and Jazayeri, 2014; Pouget et al., 2013) and the notion of 'confidence' used in metacognitive research are two different manifestations of the same concept of Bayesian probability. First, we note that 'uncertainty' and 'confidence' are merely the inverse (or reciprocal) of one another, so the choice of emphasis is not an important difference. Instead, the critical difference is that 'confidence' in the metacognitive field is a single number, such as a numerical rating, whereas 'uncertainty' in the Bayesian computation field is a property of an array of numbers, such as a distribution of firing rates across neurons*."

Meyniel, Sigman & Mainen 2015

[140] *"We propose that confidence should be used to refer to the probability that a choice is correct, which we denote p(z = k|d = k, Image, Vestib). This definition has a long history in psychophysics and has been recently used in several studies. This is also what many authors call confidence, even if they don't always formally define it as such. This definition not only applies to decisions, but also to confidence in propositions, or potentially even to aspects of self-confidence. For example, suppose you are asked to express your confidence in the following proposition: 'Nigeria is the most populous African country.' This amounts to asking your confidence in choosing this proposition versus 'Nigeria is not the most populous African country.' Thus, as for decision confidence, the confidence in this proposition can be defined as the probability that the decision, 'Nigeria is the most populous African country', is correct. The same applies to some aspects of self-confidence. Lionel Messi is presumably highly self-confident in his ability to score in soccer games because the probability that the proposition 'I will score' (as opposed to 'I will not score') is correct tends to be high. The concept that unifies all of these seemingly different types of confidence is that they are about a choice being correct, even if only hypothetically, such that confidence can be expressed probabilistically by p(z = k|d = k, evidence). Here we focus mostly on confidence about decisions, but our conclusions apply just as well to propositions. When compared to the posterior p(z|Image, Vestib) over all possible choices, confidence is the probability mass of this posterior for one particular (overt or covert) choice. But does it ever make sense to maintain a separate measure of confidence rather than continuing to use the full posterior? In other words, why would you use a limited summary, confidence, when the entire posterior distribution is available? This is because confidence is in fact the only quantity that is needed in a wide variety of tasks. It is particularly important in sequential decisions, when subsequent choices depend on previous decisions. One example of such a task is a post-decision wager, in which subjects are asked to place a bet on whether their decision was correct. The optimal size of the wager, the investment, depends on the degree of belief that the initial choice was correct, with a higher wager when confidence is high. These types of post-decision wagers can be studied in the laboratory, even in animals. One example is a time investment task, initially introduced to study confidence in rats, that requires the decision maker to first gather evidence about which of several choice options is rewarded. After a choice is made, the reward is delayed for a randomized interval and it is up to the decision maker to choose how long to wait for this reward. To not wait in vain, it only makes sense to wait extended periods if the decision maker is confident of their choice. In fact, it can be shown that there is no need to store the posterior distribution over the choices for this kind of task: the probabilities associated with the choices that the subject did not select are irrelevant, the only required quantity is the probability that the selected choice is correct. Confidence can also be important for learning from feedback (Box 2) and group decision-making. However, confidence is not always the appropriate measure to use, even in sequential decisions. For instance, if a subject receives further information relevant to a previously taken choice, then the entire posterior distribution over the latent variable z, p(z|Image, Vestib), needs to be updated in the face of new evidence. Even in this situation, confidence may be a computationally efficient summary statistic to use instead of the full posterior distribution. Consider, for example, complex environments in which the posterior distribution might require an inordinate amount of data to learn, involve extremely complex inference to compute or require large neuronal resource to store. As a result, the posterior distributions computed will only be a rough approximation of the true posterior. Using*

confidence in these situations as an approximation to the full posterior can be the computationally appropriate strategy that beats other solutions that were optimal if more information were available."

Pouget et al. 2016

[141] *"… these theories have so far been explored and tested mostly in the domain of perceptual processing (Bejjanki et al., 2011, Berkes et al., 2011, Deneve et al., 1999, Fiser et al., 2010, Kim and Basso, 2010, Ma et al., 2006). It remains an open question to what extent probabilistic computation holds beyond low-level sensory and motor representations: e.g., the belief that 'it may rain tomorrow,' a reward expectation, etc. Forming probability distributions by simulating internal models could serve as the basis for a distributional neural representation of confidence in a variety of problems. There do exist a number of models for higher-level computations, for instance involving sampling schemes with integration of samples internally generated, e.g., for evaluating general-knowledge statements (Gigerenzer et al., 1991, Juslin et al., 2007, Koriat, 2012), for learning and goal-directed decisions (Hinton and Dayan, 1996, Legenstein and Maass, 2014, Solway and Botvinick, 2012), and even for probabilistic abstract reasoning (Chater et al., 2006, Denison et al., 2013, Vul et al., 2009).*"

Meyniel, Sigman & Mainen 2015

"From this modest premise, our seemingly lofty aim is to bridge the gap between psychology on the one hand and neuroscience on the other. The foundation for our approach is first to recognize that, semantically, confidence is a property (degree, probability, etc.) that describes or modifies a referent (belief, response, memory, future event, etc.). Therefore it is impossible to refer precisely to confidence without specifying the object to which it pertains. In common usage the referent is often not made explicit and this is likely to contribute to conceptual confusion. We propose that the same general formal notion of confidence as Bayesian probability can be applied to widely different structures and processes. These include populations of neurons, neural functions, behavioral outputs, persons, etc."

Meyniel, Sigman & Mainen 2015

[142] "*To sum up, we have given different names (summary confidence, distributional confidence) to aspects of confidence that we think are worth keeping distinct. We have described how, for simple examples, summary confidence can be derived normatively from the distributional confidence information conveyed by probabilistic neural representations. We will go into more complexity later, with less direct routes and deviations from optimality (see A Brain-Scale, Hierarchical Neural Architecture for Confidence and also The Rough Edges). For the moment, the implications of this basic conceptualization can be related to the classic literature on confidence. We suggest that some confusion in the field of confidence studies is due to the conflation of distributional and summary forms. We propose that in decision-making, choice and confidence can be read out from the same neural representation (Kepecs and Mainen, 2012, Kepecs et al., 2008). This view resembles the "shared encoding" hypothesis reported by Grimaldi et al. (2015) or 'first-order model' (Timmermans et al., 2012) in which the same stream of information accounts for choice and confidence. However, these models are usually thought to entail that the same circuitry underpins choice and confidence (Grimaldi et al., 2015). We suggest the opposite: the mechanisms that read out a choice and a summary confidence from the same representation must be partly different, simply because they result in different things. Such 'parallel processing' of choice and confidence is the landmark of 'dual route models' (Timmermans et al., 2012), but our framework rejects a pure parallelism by assuming a common initial representation. Our view could therefore seem closer to 'hierarchical models' (Fleming and Dolan, 2012, Fleming et al., 2012, Timmermans et al., 2012). However, such models make a distinction between a first-order level (choice) and a second-order level (confidence) processing. This distinction is a landmark in the metacognition literature. In our view, there is no need for such a terminology: readout of choice and confidence are simply different without one being subordinate to the other."*

Meyniel, Sigman & Mainen 2015

[143] "*Interestingly, the readout of confidence can be selectively impaired in specific domains. Fleming and colleagues reported such a case: patients with brain lesions in the anterior PFC had preserved performance in the memory and perceptual domains and degraded confidence judgments specifically in the perceptual domain (Fleming et al., 2014). The fact that choice performance was preserved rules out the possibility that perception or memory, as a whole, were impaired, and points to the readout of confidence itself. This example suggests that*

---

one region alone does not suffice to read out confidence: at a minimum, it should involve a circuitry to collect specific inputs from different cognitive domains."

Meyniel, Sigman & Mainen 2015

[144] *"Dopamine neurons of the ventral tegmental area (VTA) and substantia nigra have long been identified with the processing of rewarding stimuli. These neurons send their axons to brain structures involved in motivation and goal-directed behavior, for example, the striatum, nucleus accumbens, and frontal cortex. Multiple lines of evidence support the idea that these neurons construct and distribute information about rewarding events…*

Surprisingly, after repeated pairings of visual and auditory cues followed by reward, dopamine neurons change the time of their phasic activation from just after the time of reward delivery to the time of cue onset. In one task, a naive monkey is required to touch a lever after the appearance of a small light. Before training and in the initial phases of training, most dopamine neurons show a short burst of impulses after reward delivery. After several days of training, the animal learns to reach for the lever as soon as the light is illuminated, and this behavioral change correlates with two remarkable changes in the dopamine neuron output: (i) the primary reward no longer elicits a phasic response; and (ii) the onset of the (predictive) light now causes a phasic activation in dopamine cell output. The changes in dopaminergic activity strongly resemble the transfer of an animal's appetitive behavioral reaction from the US to the CS. In trials where the reward is not delivered at the appropriate time after the onset of the light, dopamine neurons are depressed markedly below their basal firing rate exactly at the time that the reward should have occurred. This well-timed decrease in spike output shows that the expected time of reward delivery based on the occurrence of the light is also encoded in the fluctuations in dopaminergic activity (18)."

Schultz, Dayan & Montague 1997

"Dopamine neurons are therefore excellent feature detectors of the 'goodness' of environmental events relative to learned predictions about those events. They emit a positive signal (increased spike production) if an appetitive event is better than predicted, no signal (no change in spike production) if an appetitive event occurs as predicted, and a negative signal (decreased spike production) if an appetitive event is worse than predicted."

Schultz, Dayan, Montague 1997

[145] "*What has been learned must sometimes be unlearned in a changing world. Yet knowledge updating is difficult since our world is also inherently uncertain. For instance, a heatwave in winter is surprising and ambiguous: does it denote an infrequent fluctuation in normal weather or a profound change? Should I trust my current knowledge, or revise it? We propose that humans possess an accurate sense of confidence that allows them to evaluate the reliability of their knowledge, and use this information to strike the balance between prior knowledge and current evidence. Our functional MRI data suggest that a frontoparietal network implements this confidence-weighted learning algorithm, acting as a statistician that uses probabilistic information to estimate a hierarchical model of the world*."

Meyniel & Dehaene 2017

[146] "*Metacognition, the ability to internally evaluate our cognitive processes, is critical for adaptive behavioral control, particularly as many real-life decisions lack immediate feedback. Specifically, action outcomes can be ambiguous, delayed, occur only after a sequence of subsequent decisions, or might never occur at all. Yet behavioral and neural evidence indicate that subjects are able to evaluate their choices online in the absence of immediate feedback, forming estimates of decision confidence and detecting and correcting response errors*."

Rouault, Dayan & Fleming 2019

[147] "*Although both JOLs and metacomprehension judgments typically demonstrate above-chance relative accuracy, the calibration and relative accuracy of these judgments is not always impressive (Maki 1998b, Nelson & Dunlosky, 1991; Weaver & Kelemen, 1997)."*

Serra & Metcalfe 2009

[148] *"Much research on metacognition has focused on the accuracy of monitoring judgments, which has mainly been conceptualized in two ways: calibration and relative accuracy. The calibration of one's judgments refers to a difference score between the mean of one's predictive judgments and one's performance on the task being*

judged or predicted. Consider hypothetical participants in a laboratory study on metamemory (metacognition about memory). While studying paired associate items (e.g., two unrelated nouns) for the criterion test, the participants make JOLs on a 0%-100% scale indicating the percent-likelihood that they will correctly remember each item on a test. Suppose that the overall mean of their JOLs is 88%. Assuming that the participants correctly remember a mean of 66% of the items on the test, their calibration score will be +22%, indicating overconfidence. A group of participants in such a study would be said to demonstrate 'good' calibration if the overall mean of their JOLs did not significantly differ from the overall mean of their performance scores. The relative accuracy of one's judgments refers to a measure of how well one's judgments differentiate performance on the cognitive task being judged. This measure is usually calculated by computing a gamma correlation between one's judgments about individual items and performance on those same test items. Like a Pearson correlation, gamma can range from –1.0 to +1.0. If a hypothetical participant gives mostly high JOLs to items they will remember on the test and low JOLs to items they will not remember on the test, their gamma correlation will be positive. Doing the opposite will result in a negative correlation. Assigning JOLs to items at random tends to result in a correlation of zero (or a correlation might even be incalculable). Typically, a gamma correlation is calculated for each participant. The mean gamma correlation is then calculated across all of the participants in a group to estimate the relative accuracy of their judgments; if the mean is significantly greater than zero, the group's judgments are said to be above chance."

Serra & Metcalfe 2009

[149] "*We are constantly flooded with information that helps form our beliefs about reality (e.g., via the Web, advertising, colleagues, and friends). Understanding how beliefs are formed is critical, as beliefs drive our actions and decisions. Normative theories assume beliefs are adjusted according to Bayes' Rule. Indeed, this assumption often holds when people in- corporate favorable news into their existing beliefs. However, for unfavorable news people show an aversion to incorporating new information. Specifically, they discount the strength of the new information leading to noisy posterior beliefs. This tendency to selectively ignore negative information is known as the 'good news/bad news effect'. For example, people adjust their beliefs regarding their level of intelligence and physical attractiveness when they receive in- formation indicating they are more intelligent and attractive than they had assumed. However, they relatively fail to adjust their beliefs in response to information suggesting they rate lower on these attributes than they had previously thought. In addition, when learning that their risk of experiencing future negative events, such as cancer, is higher than they had expected, people are less likely to integrate these data into their prior beliefs relative to a situation when they learn that their risk is lower than expected."*

Sharot et al. 2012

[150] "*Kruger and Dunning (1999) suggested that, across many intellectual and social domains, it is the poorest performers who hold the least accurate assessments of their skill and performances, grossly overestimating how well their performances stack up against those of their peers. For example, students performing in the bottom 25% among their peers on tests of grammar, logical reasoning, and humor tended to think that they are performing above the 60th percentile (Kruger & Dunning, 1999). Further, this pattern has been conceptually replicated among undergraduates completing a classroom exam (Dunning, Johnson, Ehrlinger, & Kruger, 2003), medical students assessing their interviewing skills (Hodges, Regehr, & Martin, 2001) clerks evaluating their performance (Edwards, Kellner, Sistrom, & Magyari, 2003), and medical lab technicians evaluating their on-the-job expertise (Haun, Zeringue, Leach, & Foley, 2000)."*

Ehrlinger et al. 2008

[151] "*In essence, we argue that the skills that engender competence in a particular domain are often the very same skills necessary to evaluate competence in that domain—one's own or anyone else's. Because of this, incompetent individuals lack what cognitive psychologists variously term metacognition (Everson & Tobias, 1998), metamemory (Klin, Guizman, & Levine, 1997), metacomprehension (Maki, Jonas, & Kallod, 1994), or self-monitorings kills (Chi, Glaser, & Rees, 1982). These terms refer to the ability to know how well one is performing, when one is likely to be accurate in judgment, and when one is likely to be in error. For example, consider the ability to write grammatical English. The skills that enable one to construct a grammatical sentence are the same skills necessary to recognize a grammatical sentence, and thus are the same skills necessary to determine if a grammatical mistake has been made. In short, the same knowledge that underlies the ability to produce correct judgment is also the knowledge."*

Kruger & Dunning 1999

[152] *"Several lines of research are consistent with the notion that incompetent individuals lack the metacognitive skills necessary for accurate self-assessment. Work on the nature of expertise, for instance, has revealed that novices possess poorer metacognitive skills than do experts. In physics, novices are less accurate than experts in judging the difficulty of physics problems (Chi et al., 1982). In chess, novices are less calibrated than experts about how many times they need to see a given chessboard position before they are able to reproduce it correctly (Chi, 1978). In tennis, novices are less likely than experts to successfully gauge whether specific play attempts were successful (McPherson & Thomas, 1989). These findings suggest that unaccomplished individuals do not possess the degree of metacognitive skills necessary for accurate self-assessment that their more accomplished counterparts possess. However, none of this research has examined whether metacognitive deficiencies translate into inflated self-assessments or whether the relatively incompetent (novices) are systematically more miscalibrated about their ability than are the competent (experts)."*

Kruger & Dunning 1999

[153] "*We report two studies in which participants completed a popular performance-based measure of analytic-thinking disposition, the CRT (Frederick, 2005), and were subsequently asked to estimate how many of the items they had gotten correct (Mata, Ferreira, & Sherman, 2013; Noori, 2016). Following Kruger and Dunning (1999), we hypothesized that participants who performed poorly on the CRT would overestimate their performance to a greater extent than would those who performed well (i.e., less-analytic people should be more poorly calibrated). In addition, participants were also asked to self-report their need or desire to think analytically using the NC scale. We predicted a Dunning–Kruger effect, such that participants who performed particularly poorly on the CRT (indicating an intuitive or non-analytic-thinking disposition) would overreport the degree to which they were disposed to analytic thinking. In Study 2, we used an independent assessment of analytic thinking—the heuristics-and-biases inventory (Toplak, West, & Stanovich, 2011, 2014)—to assess whether nonanalytic individuals are genuinely worse at recognizing their bias. Put differently, participants who were decidedly nonanalytic based on the performance measure should be less-suited to assess their degree of analyticity on a self-report measure, leading to poor calibration in terms of both estimated CRT accuracy and self-reported NC. Our results provide empirical support for Dunning–Kruger effects in both estimates of reasoning performance and self- reported thinking disposition. Particularly intuitive individuals greatly overestimated their performance on the CRT—a tendency that diminished and eventually reversed among increasingly analytic individuals. Moreover, self-reported analytic- thinking disposition—as measured by the Ability and Engagement subscales of the NC scale—was just as strongly (if not more strongly) correlated with estimated CRT performance than with actual CRT performance. In addition, an analysis using an additional performance-based measure of analytic thinking—the heuristics-and-biases battery—revealed a systematic miscalibration of self-reported NC, where- in relatively intuitive individuals report that they are more analytic than is justified by their objective performance. Together, these findings indicate that participants who are low in analytic thinking (so-called intuitive thinkers) are at least somewhat unaware of (or unresponsive to) their propensity to rely on intuition in lieu of analytic thought during decision-making. This conclusion is consistent with previous research that has suggested that the propensity to think analytically facilitates metacognitive monitoring during reasoning (Pennycook et al., 2015b; Thompson & Johnson, 2014). Those who are genuinely analytic are aware of the strengths and weaknesses of their reasoning, whereas those who are genuinely nonanalytic are perhaps best described as 'happy fools' (De Neys et al., 2013)."*

Pennycook et al. 2017

[154] "*The present study focuses on this critical error sensitivity issue. Preschool children were given a classic version of a number conservation task in which an intuitively cued response conflicted with the correct conservation response and a control version in which this conflict was not present. After solving each version children were asked to indicate their response confidence. Results showed that in contrast with children who gave a correct conservation response, preschoolers who erred showed a sharp confidence decrease after solving the classic conflict problem. This suggests that non-conserving preschoolers detect that their response is questionable and are less ignorant about conservation than their well-documented errors might have previously suggested."*

De Neys, Lubin & Houdé 2014

155 "*This contraposition indicates that the neurological (e.g., De Neys et al., 2008) and physiological (De Neys et al., 2010) conflict detection signals may be relatively effective, but the response to this signal may actually be rather ineffective. De Neys et al. (2013) found a large (15%) decrease in confidence for the bat-and-ball problem relative to a control, but 82% confidence is still quite high."*

Pennycock et al. 2017

156 "*Although it is clear that people are often biased, the nature of this bias is poorly understood. Some authors claim that people reason heuristically by default and that most of the time they are simply not aware that their intuitions might be wrong. The dominance of intuitive thinking is attributed to a failure to monitor the output of the heuristic reasoning process. In this view, because of lax monitoring, people fail to detect that an intuitive response conflicts with the response favored by probability. The problem is that people do not know that their judgment is biased. This view has been popularized by the work of authors such as Kahneman (2002) and Evans (1984, 2003).*"

De Neys, Vartanian & Goel (2008)

157 *"Intuitive or lay theories are thought to influence almost every facet of everyday cognition. People appeal to explanatory relations to guide their inferences in categorization, diagnosis, induction, and many other cognitive tasks, and across such diverse areas as biology, physical mechanics, and psychology (Gopnik & Wellman, 1994; Keil, 1998; Murphy & Medin, 1985; Murphy, 2000). Individuals will, for example, discount high correlations that do not conform to an intuitive causal model but overemphasize weak correlations that do (Chapman & Chapman, 1969). Theories seem to tell us what features to emphasize in learning new concepts as well as highlighting the relevant dimensions of similarity (Murphy, 2002). Intuitive theories have also been heavily emphasized in accounts of the cognitive development of children (Gelman & Koenig, 2002) and even of infants (Spelke, Breinliinger, Macomber, & Jacobson, 1992). Concepts seem to be embedded within larger sets of explanatory relations that are essential to understanding the structure of the concepts themselves, how they are learned, and how they change over time. But even as theories have become more central to the study of concepts, it is also now evident that folk theories are rarely complete or exhaustive explanations in a domain (Wilson & Keil, 1998). Indeed, even the theories used daily to guide scientific research are now considered to be incomplete, or at least less formally logical than classical views assumed them to be (Boyd, 1991; Salmon, 1989, 1998). Science-in-practice is often driven by hunches and vague impressions."*

Rozenblit & Keil 2002

158 *"The pattern of results so far indicates a special difficulty in calibrating one's explanatory knowledge about devices. In the studies with procedures and movies, the participants were well calibrated. In the study with factual knowledge about capitals, the participants were overconfident but markedly less so than with explanatory knowledge of devices… To summarize, studies with devices and natural phenomena both show large drops in knowledge estimates. Procedures and Narratives show no drop, while Geography Facts show only a small drop. The results demonstrate large differences in knowledge calibration across knowledge domains, casting serious doubt on the meaningfulness of 'general overconfidence' about knowledge. The studies also raise intriguing possibilities about the mechanism behind over-confidence, which we address in the next few studies…One conclusion that can be drawn from this research is that the well-established blanket approach to overconfidence with 'general knowledge' is almost certainly misleading. Large inter-domain differences in calibration imply that structural properties of knowledge have a powerful impact on the process of knowledge assessment. 'General knowledge' is a chimera— a mythological composite creature. Taking it seriously distracts from interesting questions about how knowledge assessment works, and the theoretically important issues of how the structural properties of knowledge influence calibration."*

Rozenblit & Keil 2002

159 "*To clarify the distinctive nature of our proposal it is useful to briefly consider prior research on overconfidence. Relevant research in the judgment and decision-making tradition has used the disparity between people's average confidence levels for their answers to almanac questions and the proportion of correct answers to argue that people are overconfident (Fischhoff, 1982; Lichtenstein & Fischhoff, 1977; Yates, Lee, & Shinotsuka, 1996; Yates, Lee, & Bush, 1997). This tradition, however, does not focus on how illusions of knowing might differ across kinds of knowledge. Lumping diverse kinds of knowledge into a hypothetical 'general knowledge' and looking for an overall overconfidence effects may well obscure large differences in*

calibration between knowledge types. The cognitive psychology literature on text comprehension also suggests overconfidence about one's knowledge. People are often poor at detecting when they have failed to understand a piece of text, both as adults (Glenberg & Epstein, 1985; Glenberg, Wilkinson, & Epstein, 1982; Lin & Zabrucky, 1998) and as children (Markman, 1977; Markman, 1979). In contrast, the current studies are concerned with people's ability to assess the knowledge they have before coming into the lab, rather than things learned in the course of an experiment. The implications of our research are different: they tell us less about how people learn when reading, and more about individuals' intuitive theories about how knowledge is stored and about the mismatch between what they think they already know and what they really know. Another area of research has focused on meta-cognition and feelings of knowing (FOK) (Koriat, 1995; Metcalfe, Schwartz, & Joaquim, 1993). One recent analysis considers the two main models for FOK to be the cue familiarity model and the accessibility model (Koriat & Levy-Sadot, 2001). The accessibility model claims that the ease of accessing information prompted by the target drives FOKs. The cue familiarity model claims that FOK judgments are elicited by the familiarity of the cues themselves…

Overconfidence also exists in areas that have little to do with knowledge. Participants have been shown to be overconfident about their future performance on motor tasks (e.g., West & Stanovich, 1997), their abilities compared to other people's abilities (e.g., Kruger & Dunning, 1999), and about their competence to perform a broad range of tasks (Bjork, 1998)."

Rozenblit & Keil 2002

"Folk theories, we claim, are even more fragmentary and skeletal, but laypeople, unlike some scientists, usually remain unaware of the incompleteness of their theories (Ahn & Kalish, 2000; Dunbar, 1995; diSessa, 1983). Laypeople rarely have to offer full explanations for most of the phenomena that they think they understand. Unlike many teachers, writers, and other professional 'explainers,' laypeople rarely have cause to doubt their naive intuitions. They believe that they can explain the world they live in fairly well. They are novices in two respects. First, they are novice 'scientists'—their knowledge of most phenomena is not very deep. Second, they are novice epistemologists—their sense of the properties of knowledge itself (including how it is stored) is poor and potentially misleading…. We argue here that people's limited knowledge and their misleading intuitive epistemology combine to create an illusion of explanatory depth (IOED). Most people feel they understand the world with far greater detail, coherence, and depth than they really do. The illusion for explanatory knowledge–knowledge that involves complex causal patterns—is separate from, and additive with, people's general overconfidence about their knowledge and skills. We therefore propose that knowledge of complex causal relations is particularly susceptible to illusions of understanding."

Rozenblit & Keil 2002

[160] *"People's estimations of the future are often unrealistically optimistic. A problem that has puzzled scientists for decades is why human optimism is so pervasive when reality continuously confronts us with information that challenges these biased beliefs. According to influential learning theories, agents should adjust their expectations when faced with disconfirming information. However, this normative account is challenged by observations that providing people with evidence that disconfirms their positive outlook often fails to engender realistic expectations. For example, highlighting previously unknown risk factors for diseases is surprisingly ineffective at altering an individual's optimistic perception of their medical vulnerability. Even experts show worrying optimistic biases. For instance, financial analysts expect improbably high profits and family law attorneys underestimate the negative consequences of divorce. The wider societal importance of these errors derives from the fact that they reduce precautionary actions, such as practicing safe sex or saving for retirement, On the upside, optimistic expectations can lower stress and anxiety, thereby promoting health and well-being. Although the existence of unrealistic optimism has been extensively documented the biological and computational principles that help to maintain optimistically biased predictions in the face of reality are unknown. Notably, such biases are not explained by theories assuming equal learning across outcome valence."*

Sharot, Korn & Dolan 2011

[161] *"What happens in the special case where the learner has no prior experience, as is the case with young children? In this case, they are over-confident, i.e. they have little or no doubt that they will succeed, nor do they expect that major and sustained efforts will be required. It is generally accepted that the systematic overconfidence of advanced beginners has an adaptive function, enabling them to engage more readily in new*

cognitive tasks. Since overconfidence is not confirmed by the results, it is followed by a phase of underconfidence, which jeopardises the motivation to persist in learning. »

Proust 2019

[162] *"In many situations, the low predictability of these judgments can be attributed to participants' basing them on information that is not diagnostic of their future performance. As described by Metcalfe, Schwartz, and Joaquin (1993):*

> *[When] making judgments related to the external world rather than to their internal states and abilities, people use other heuristics or rules of thumb. These heuristics approximate the uncertain quantity indirectly, rather than measuring the quantity itself. Because they do not measure the quantity directly, such heuristics may result in biases and errors. (p. 860).*

In other situations (e.g., delayed JOLs, as we discuss later), calibration and relative accuracy can be quite good because participants base their judgments on information that is diagnostic of their future performance. This section reviews some heuristics that are sometimes used to inform metacognitive judgments and that can produce errors and illusions in monitoring. This section is not meant to describe all of the potential errors and poor heuristics that might arise in metacognitive monitoring, but simply to provide some illustrative examples. It also suggests some ways that the errors associated with these particular cues might materialize in the classroom. As we learn more about heuristics that lead to accurate and inaccurate metacognitive judgments, we can better understand which heuristics aid or hinder metacognitive accuracy."

Serra & Metcalfe 2009

[163] "*Learners' familiarity with the information being judged can have an influence on their metacognitive judgments. Because this experience of familiarity often results from prior exposures to and learning of the information, it can be diagnostic of a greater likelihood that the information is known (i.e., it often results in accurate judgments). Unfortunately, familiarity can also arise in situations where it is unrepresentative of knowledge. For example, participants in a study by Reder and Ritter (1992) solved difficult arithmetic problems. After each was presented, they had to quickly choose whether to calculate the answer to each problem or recall it from memory (calculation, of course, was the only option the first time a problem was presented). Participants received 50 points for correctly recalling an answer and 5 points for correctly calculating an answer (but only if the selections and responses were made within the designated time limits). Participants were later paid .05 cents for each point earned. By manipulating the occurrence of specific numbers in the set of problems, Reder and Ritter manipulated the participants' familiarity for the numbers present in the problems independent of their memory for specific problems and their solutions. Participants were able to use their familiarity with the numbers—both independently and as whole problems—to quickly decide whether they knew the answers to the problems. This strategy proved helpful when the specific problem had actually been presented in accordance with the participants' familiarity for the numbers, but proved to be faulty when familiar numbers were combined into novel problems—problems for which the participants could not actually recall an answer*."

Serra & Metcalfe 2009

[164] *"When participants are required to list the reasons for their choice of an answer to almanac questions (Koriat et al., 1980, 2008), they typically mention logical or rational considerations. However, it is clear that some of the representations that tip the balance in favor of one choice or the other consist of associations and images that cannot be expressed in a propositional form, and some operate bellow full consciousness. Indeed, studies of the illusory-truth effect indicate that the mere familiarity and fluency of a statement that are caused by its repetition or by its context can influence the perceived truth of that statement (Arkes, Hackett, & Boehm, 1989; Bacon, 1979; Hasher, Goldstein, & Toppino, 1977; Unkelbach & Stahl, 2009)."*

Koriat 2012

[165] *"The overconfidence bias is assumed to follow from the basic assumption that confidence judgments rely on reliability as a cue for validity. Reliance on reliability—the consistency with which a choice is supported—may instill inflated confidence because reliability is practically always higher than validity. Indeed, although confidence judgments yielded an overconfidence bias when evaluated against correctness, these judgments were not markedly inflated when evaluated against several indexes of self- consistency. It should be stressed that the SCM account of overconfidence does not postulate a specific bias like the biases pro- posed in previous*

discussions (Koriat et al., 1980; Ronis & Yates, 1987; Taylor & Brown, 1988). Rather, the bias is assumed to stem from the very basis of confidence judgments."

Koriat 2012

166 *"Current-knowledge heuristic. Sometimes, after information has been obtained or understood, people think they knew or understood it all along. Participants in a study by Fischhoff (1975) read passages detailing outcomes associated with historical events (such as a battle). They then judged if they would have predicted the outcome before reading the passage. Fischhoff demonstrated that these participants could not avoid using their knowledge of the outcome when making this judgment; they even judged that they would have correctly predicted highly unlikely events (some of which were false). The participants in his study demonstrated hindsight bias—a tendency to use new knowledge when thinking about the past—without even knowing that they were doing so."*

"Association heuristics. Some information seems easier to understand or remember when it is studied than it will actually be to remember or apply later on a test. Koriat and Bjork (2006) termed such an illusion foresight bias and demonstrated a type of paired-associate that produces such an effect. These pairs were composed of two words that had a strong backwards association but a weak forward association (i.e., one would be likely to think of the first word when shown the second word but not likely to think of the second word when shown the first word). For example, consider the pair 'fire—blaze'. The word 'blaze' is almost always freely-associated to the word 'fire', but 'fire' rarely—if ever—is freely-associated to the word 'blaze'. When such pairs are studied and judged in a typical metamemory procedure, the presence of both words at study makes them seem highly related. At test, however, the stimulus word (fire) is not actually likely to produce the response word (blaze). The association strength present at study produces the illusion that the response word will easily be recalled at test."

"Heuristics that can cause illusions of knowing. One's experience with learning materials sometimes causes the illusion that the materials have been understood when in fact they have not. Participants in a study by Glenberg, Wilkinson, and Epstein (1982) demonstrated an illusion of knowing (i.e., their judgments were overconfident) when asked to rate their comprehension for texts containing factual contradictions. Participants often failed to find these contradictions yet rated their understanding of the texts as being high. This even occurred when factual contradictions were in two adjacent sentences (Glenberg et al., 1982). These findings suggest that readers do not attempt to monitor their understanding across a whole text, but rather at lower levels such as at the per-sentence level."

Serra & Metcalfe 2009

167 "*As described in the previous section of this chapter, metacognitive judgments are prone to errors, biases, and metacognitive illusions. Experience with and information about these illusions can help to reduce some of these biases. Koriat and Bjork (2006) described one such illusion—foresight bias—in which some to-be-studied pairs of words have a strong backwards association but a weak forward association (e.g., fire—blaze). The presence of both words at study but not at test produced overconfident JOLs for these items. As in King, Zechmeister, and Shaughnessy (1980), Koriat and Bjork (2006) demonstrated that study-test practice reduced this bias, but it did not transfer to new items. Explicit training about the foresight bias and the type of item that causes it, however, not only reduced the bias but also transferred to new items."*

Serra & Metcalfe 2009

168 *"Most theories predict that when people indicate that they are highly confident they are producing their strongest responses. Hence, if such a high confidence response is in error it should be overwritten only with great difficulty. In contrast to this prediction, we have found that people easily correct erroneous responses to general information questions endorsed as correct with high- confidence, so long as the correct answer is given as feedback. Three potential explanations for this unexpected hypercorrection effect are summarized. The explanation that is tested here, in two experiments, is that after a person commits a high-confidence error the correct answer feedback, being surprising or unexpected, is given more attention than is accorded to the feedback to low-confidence errors. This enhanced attentional capture leads to better memory. In both experiments, a tone detection task was presented concurrently with the corrective feedback to assess the attentional capture of feedback stimuli. In both, tone detection was selectively impaired during the feedback to high confidence errors. It was also negatively related to final performance, indicating that the attention not*

devoted to the tone detection was effectively engaged by the corrective feedback. These data support the attentional explanation of the high-confidence hypercorrection effect."

Butterfield & Metcalfe 2006

[169] *"The purpose of the current research is to reexamine the effect of feedback on retention of initially correct responses. Of course, we are not arguing against the fact that correcting memory errors is a key purpose of feedback. Instead, we believe that feedback also functions as an error-correction mechanism for correct responses, albeit for a different type of error. When individuals make a correct response but are not confident in the response, there is a discrepancy between the subjective and objective correctness of their answers. In other words, low-confidence correct responses reflect an error of metacognitive monitoring, which in this context refers to the ability to assess the accuracy of one's own performance on a test (Barnes, Nelson, Dunlosky, Mazzoni, & Narens, 1999; Koriat & Goldsmith, 1996; Nelson & Narens, 1990). Feedback that confirms the correctness of low-confidence responses should enable learners to reduce the discrepancy between their perceived and actual performance by allowing them to adjust their subjective assessments of their knowledge. Further, if feedback allows learners to correct initial metacognitive errors, then it should also enhance long-term retention of the correct responses and improve the accuracy of metacognitive monitoring on subsequent tests. Thus, our hypothesis in this research was that, just as feedback helps correct memory errors, feedback will also help correct meta-cognitive errors and will improve retention of low-confidence correct responses…*

Our hypothesis is that feedback serves to correct the metacognitive error inherent in low-confidence correct responses, much like it does for high-confidence errors in the hypercorrection effect. However, we believe that the correction of these two types of metacognitive error probably leads to better retention through different mechanisms. As described above, retention may be enhanced following high-confidence errors because a feeling of surprise causes subjects to pay more attention to feedback (Butterfield & Metcalfe, 2006). In contrast, we think that providing feedback after low-confidence correct responses might enhance retention by enabling learners to strengthen the association between the cue and response and to inhibit any competing responses…

the current experiments provide clear evidence that low-confidence correct responses do benefit from feedback and that feedback improves students' metacognitive judgments about their knowledge. Taken together, the two novel findings support the idea that a low-confidence correct response represents an error in metacognitive monitoring that can be corrected through feedback. Providing feedback after low-confidence correct responses enables learners to eliminate the discrepancy between perceived and actual correctness of the response. Feedback after both correct and incorrect responses on tests is a critical aspect of learning."

Butler, Karpicke & Roediger 2008

[170] "*Although metacognitive abilities are often treated as stable characteristics of individuals (Allen et al., 2017; Fleming, Weil, Nagy, Dolan, & Rees, 2010; McCurdy et al., 2013; Song et al., 2011), several lines of research hint at their malleability. For instance, practicing meditation boosts the accuracy of retrospective confidence judgments about recognition memory decisions (Baird, Mrazek, Phil- lips, & Schooler, 2014) and monitoring of decision errors can be modulated by drugs (Hester et al., 2012; Hauser et al., 2017) and brain stimulation (Harty et al., 2014). Moreover, recent work has identified distinct neural substrates in the frontal and parietal lobes supporting metacognitive monitoring across a range of tasks (Al- len et al., 2017; Baird, Smallwood, Gorgolewski, & Margulies, 2013; Cortese, Amano, Koizumi, Kawato, & Lau, 2016; Fleming et al., 2010; McCurdy et al., 2013; see Fleming & Dolan, 2012, for a review), suggesting the potential for targeted modulation of metacognition independently of changes in first-order performance. Previous attempts to improve metacognitive ability (confidence calibration) through explicit instruction, practice, feedback, or a combination of these manipulations have led to mixed results, with some studies documenting increases, and others documenting null findings (e.g., Adams & Adams, 1958; Bol, Hacker, O'Shea, & Allen, 2005; Lichtenstein & Fischhoff, 1980; Nietfeld & Schraw, 2002; Renner & Renner, 2001; Sharp, Cutler, & Penrod, 1988). One potential explanation for such heterogeneity of results is that training may impact first-order performance, thus masking subtle changes in metacognition because they are positively correlated (Fleming & Lau, 2014; Sharp et al., 1988)."*

Carpenter et al. 2019

171 "*…it remains poorly understood as to whether metacognition relies on a domain-general resource that is 'applied' to the task at hand, or whether different metacognitive processes are engaged when evaluating performance in different domains."*

Rouault et al. 2018

172 *"That we as humans do make errors in thinking, judgment, and memory is undisputed. In fact, there is a plethora of phenomena showing that we deviate in our thinking, judgment, and memory from the objective and arguably 'correct' standard'."*

Pohl 2017

173 *"Most researchers in social cognition favor the view that heuristics are simple, efficient shortcuts applied in judgment and decision-making when people face overly complex tasks, have limited time or cognitive ability, or deal with incomplete information in the world. In this light, heuristics work well in many instances, but are prone to break down in systematic ways-and whenever they do, more 'evidence' has been found that the mind is flawed in its reasoning abilities. The traditional treatment of heuristics has been largely dominated by researchers working within the heuristics and biases program (Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1974), who have argued that human judgment often substantially deviates from optimality predictions or normative standards of logic."*

Haselton et al. 2009

174 "*According to Gigerenzer, the collection of biases has focused too much on the few faulty cases of judgment and decision-making thereby ignoring the majority of heuristics where heuristics typically lead to correct or at least useful decisions. This one-sided view may have led some researchers to conclude that 'mental illusions should be considered the rule rather than the exception' (Thaler 1991) and that 'mistakes of reason rule our mind' (Piattelli-Palmarini 1994). But this view of human rationality appears overly pessimistic…*

Gigerenzer (1991) has also asserted that non everything that looked like a cognitive illusion really is one. More specifically he argued that one could make cognitive illusions disappear by (a) avoiding too narrow norms to evaluate human performance, by (b) using an adequate statistical format, and by © using representative (instead of selected) samples of items. …

Using adequate statistical formats and representative sampling, Gigerenzer (1991; see also Gigerenzer & Hoffrage 1995) accordingly showed that some illusions could be substantially reduced or even eliminated, namely, the conjunction fallacy, … baserate neglect, … and overconfidence.

Pohl 2017

175 *"A fundamental criticism of the heuristics and biases program is that researchers might be neglecting the structure of the world in which the decision-making takes place. Ecologically-minded scientists have argued that in order to understand the mind's true cognitive abilities one needs to consider the environment in which it operates-or was designed to operate by natural selection. For example, Egon Brunswik (1955) emphasized that psychologists should study how the mind makes inferences based on the informational cues present in the natural environment, and Roger Shepard (2001) saw the mind as a mirror reflecting regularities of the physical world (see Todd & Gigerenzer, 2007). Consequently, what we call a good or a bad decision (or rational and irrational behavior) has to be judged with regard to specific decision environments rather than in a vacuum (Gigerenzer, Todd,& the ABC Research Group, 1999)…*

A related criticism addresses how much information is usually available in these decision environments and if the clear standard for comparing decision outcomes-the supposedly optimal way of thinking-should necessarily be informed by abstract standards of probability, logic, and mathematical optimization. Many traditional models of rational choice assume that humans (and animals) make inferences about the world virtually as if they were supernatural beings that have unlimited reasoning power, boundless knowledge, and unlimited time to make their decisions. However, real-world decision environments, both current and past, do not look like this and it is unrealistic to compare the human capacity for judgment and choice against such optimality predictions and assumptions (see Gigerenzer et al., 1999). For example, humans almost never have access to all of the pertinent information needed for making a decision about which mate to choose, what foods are best to eat, or which house to buy. Rather than following models of unbounded rationality, researchers pointed out that many *decisions are made in a boundedly rational way (i.e., under conditions of limited time, information, and cognitive processing) and that heuristics are psychologically plausible solutions in situations where the one best solution does not exist or cannot be reached anyhow."*

Haselton et al. 2009

176 *"From an evolutionary perspective, however, it would be surprising if the* mind were really so woefully *muddled. The mind is an intricate, evolved machine that has allowed humans to inhabit and exploit an incredible range of environments. Humans effectively solve a variety of social-ecological problems including large-scale cooperation, social exchange, habitat formation, agriculture, and cumulative culture. We are a remarkably intelligent species, capable of surviving and reproducing in a complicated and ever-changing world. Could it really be that the human mind is as deeply flawed as the literature suggests?"*

Haselton et al 2009

177 "*The term 'cognitive illusion' has evolved in analogy to the better-known domain of 'optical illusions' (see Roediger 1996). The first and main feature of a phenomenon to count as an illusion thus is that it leads to a perception, judgment or memory that reliably deviates from 'reality'. IN cases of optical and memory illusions, it may be immediately evident what constitutes reality (because subjective perception and recall can be compared to external and original stimuli, respectively), but in thinking and judgment, the matter is less clear (Gigerenzer 1996). The problem concerns how to define an objectively 'correct' judgment or decision…*

As a second criterion, the observed phenomenon needs to deviate from the normative standard in a systematic fashion (i.e. in a predictable direction) rather than just randomly. Therefore, most designs include a control group, assuming that any deviations in the control group's data result from random error alone, while the experimental group shows in addition a systematic effect…

the mechanisms eventually leading to cognitive illusions typically include a number of probabilistic processes so that an illusion will not necessarily be observed on each and every single trial, but may only become evident as a systematic bias if the data are summed across a large number of trials or participants. A third aspect of illusions is that they appear involuntarily, that is, without specific instructions or deliberate will. They just happen. ... This does not mean that motivational factors or conscious metacognitions may not be influential, too, but they are not the ultimate cause of the illusion itself. They only moderate its size…

Another aspect is that persons who have fallen prey to a cognitive illusion usually don't realize what has happened. … That is, illusioned persons are still convinced to have judged, decided, or recalled something to the best of their knowledge. As a consequence, and this constitutes the fourth cornerstone of the proposed definition, and illusion is hard if not impossible to avoid. While this is probably true for optical illusions, the criterion is much weaker for cognitive ones. For some illusions, proper instruction, careful selection of the material, or other procedural variations may reduce or even eliminate the illusion (as an example see Gigerenzer, Hertwig, Hoffrage & Sedlmeier 2008), while for other illusions, most (if not all) attempts to overcome the effect have failed (as an example, see Pohl & Hell 1996…). And finally, as the fifth point to consider and to distinguish cognitive illusions from other forms of typical errors, misunderstandings or faulty memories, illusions often appear as rather distinct from the normal course of information processing. An illusion somehow 'sticks out' as something special that 'piques our curiosity' (as Roediger 19996 put it) and thus attracts researchers to explain this unexpected but robust finding. In other words, ordinary forms of forgetting (leading to commission errors), or deviations resulting from simple misunderstandings would not be considered 'illusions'. ... This is not to say that an illusion cannot be explained with ordinary and general mechanisms of information processing. In fact, one of the theoretical goals of research in cognitive illusions is to avoid the assumption of any special mechanisms that is responsible only for this one phenomenon, but instead to explain the observed effects with what one already knows about cognitive processes in general."

Pohl 2017

178 *"On the surface, cognitive biases appear to be somewhat puzzling when viewed through an evolutionary lens. Because they depart from standards of logic and accuracy, they appear to be design flaws instead of examples of good engineering. Cognitive traits can be evaluated according to any number of performance criteria—logical sufficiency, accuracy, speed of processing, and so on. The value of a criterion depends on the question the scientist is asking. To the evolutionary psychologist, however, the evaluative task is not whether the cognitive feature is accurate or logical, but rather how well it solves a particular problem, and how solving this*

problem contributed to fitness ancestrally. Viewed in this way, if a cognitive bias positively impacted fitness, it is not a design flaw—it is a design feature."

Haselton, Nettle & Murray 2015

[179] "*On the surface, cognitive biases appear to be somewhat puzzling when viewed through an evolutionary lens. Because they depart from standards of logic and accuracy, they appear to be design flaws instead of examples of good engineering. Cognitive traits can be evaluated according to any number of performance criteria—logical sufficiency, accuracy, speed of processing, and so on. The value of a criterion depends on the question the scientist is asking. To the evolutionary psychologist, however, the evaluative task is not whether the cognitive feature is accurate or logical, but rather how well it solves a particular problem, and how solving this problem contributed to fitness ancestrally. Viewed in this way, if a cognitive bias positively impacted fitness, it is not a design flaw—it is a design feature."*

Haselton, Nettle & Murray 2015

[180] "*By cognitive bias, we mean cases in which human cognition reliably produces representations that are systematically distorted compared to some aspect of objective reality." ...*

"*An evolutionary psychological perspective predicts that the mind is equipped with function-specific mechanisms adapted for special purposes—mechanisms with special design for solving problems such as mating, which are separate, at least in part, from those involved in solving problems of food choice, predator avoidance, and social exchange (e.g., Kenrick, Neuberg, Griskevicius, Becker, & Schaller, 2010). In the evaluation of cognitive biases, demonstrating domain specificity in solving a particular problem is a part of building a case that the trait has been shaped by selection to perform that function*." *...*

"Some design features that appear to be flaws when viewed in one way are revealed to be adaptations when viewed differently. If one were to only consider the idea that selection favors the maximization of direct reproductive success, for example, the fact that human females lose reproductive capability many years before death would appear a design flaw. However, there is evidence that women in traditional societies can enhance their inclusive fitness by transferring investment to their daughters' daughters as soon as the latter are of reproductive age (Voland & Beise, 2002). Viewed in this light, female menopause might be very well designed (Hawkes, 2003). In sum, there may be many evolutionary reasons for apparent design flaws, and a close examination often provides insight into the evolutionary forces that shaped them and their functions."

Haselton, Nettle & Murray 2015

[181] "*Knowledge of biases and illusions is of course valuable. For example, demonstrating that a bias may occur in some situations but not others (context effects), or with certain classes of information and not others (content effects), can reveal structural features of the mind. Additionally, knowledge of biases and illusions may have important practical utility by preventing undesirable outcomes.*"

Haselton et al 2009

[182] "*First, selection may favor useful shortcuts that tend to work in most circumstances, though they fall short of some normative standards (heuristics); second, apparent biases can arise if the task at hand is not one for which the mind is designed (artifacts); and third, biases can arise if biased response patterns to adaptive problems resulted in lower error costs than unbiased response patterns (error management biases).*"

Haselton, Nettle & Murray 2015

[183] "*Tversky and Kahneman attributed these and other biases to the operation of mental shortcuts : 'People rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations' (1974, p. 1124). The gambler's fallacy and the conjunction fallacy are attributed to one of the most commonly invoked heuristics, representativeness, or the way in which A resembles or is representative of B. According to this account, alternating heads and tails are more representative of randomness than are series containing runs.*" *…*

"*Overall, there is ample evidence of cognitive bias and error in humans. Some of these biases may result from the use of shortcuts, which are often effective. For these effects, however, it is important to note that a* *'processing limitations' explanation is not complete. Of all possible equally economical cognitive shortcuts, why were these particular ones favored by selection?"*

Haselton, Nettle & Murray 2015

184 "*One type of artifact arises from evolutionarily novel problem formats. Gigerenzer (1997) proposed that tasks intended to assess human statistical prediction should present information in frequency (rather than probability) format, given that natural frequencies, such as the number of times an event has occurred in a given time period, are more readily observable in nature. In contrast, probabilities (in the sense of a number between 0 and 1) are mathematical abstractions beyond sensory input data, and information about the base rates of occurrence is lost when probabilities are computed (Cosmides & Tooby, 1996). Bayesian calculations involving frequencies are therefore computationally simpler than equivalent calculations involving probabilities, relative frequencies, or percentages. Whereas probability calculations need to reintroduce information about base rates, frequency calculations do not since this part of the computation is already 'done' within the frequency representation itself (Hoffrage, Lindsey, Hertwig, & Gigerenzer, 2001) …*

"A second artifact can arise from evolutionarily novel problem content. The perspective on cognitive design we have described suggests that researchers should not necessarily expect good performance in tasks involving abstract rules of logic. Falsification- based logic is sufficiently difficult for humans that university courses in logic, statistics, and research design attempt to teach it to students (with only mixed success)…

Haselton, Nettle & Murray 2015

185 "*Within this framework, many ostensible faults in human judgment and evaluation may reflect the operation of mechanisms designed to make inexpensive, frequent errors rather than occasional disastrous ones (Haselton & Nettle, 2006; Johnson et al., 2013) ...*

"*Food Aversions Lasting aversion to a food is reliably acquired, in humans and other species, following a single incidence of sickness after ingestion of the food (Garcia, Hankins, & Rusiniak, 1976; Rozin & Kalat, 1971). Given one data point (sickness following the food type on one occasion), the system treats the food as if it is always illness-inducing. There are again two possible errors here. The false positive may be inconvenient, but the false negative is more likely to be fatal. The system appears biased toward over-responsiveness to avoid illness...*

"*The error management account is similar to that for food aversions: The false negative (failing to avoid someone with a contagious disease) is highly costly, whereas the false positive (avoiding contact with a non contagious person) may have small social or interpersonal costs, but is unlikely to have significant negative fitness consequences*."

Haselton, Nettle & Murray 2015

186 "*We ultimately conclude that the mind is best described as adaptively rational. By adaptably rational we mean that the mind shows evidence of psychological design for coping with recurrent adaptive problems our ancestors encountered over evolutionary history - the mind is equipped with mechanisms that are constrained and sometimes imprecise, but nevertheless clear products of natural selection showing evidence of good design. This definition runs in contrast to the often implicit definitions of rationality used by many social science researchers, including that the mind should maximize 'accuracy,' happiness, well-being, financial return, or adherence to abstract rules of logic. We do not deny that it is useful to compare human performance to these standards, as they may be those we wish to maximize in the modem world. Instead, we challenge the idea that deviations of performance from the standards means that the human mind is deeply flawed or poorly designed*."

Haselton et al 2009

187 "*The default-interventionist model and the corresponding bias blind spot and corrective assumptions have had far-reaching impact on theorizing in the various fields that have adopted dual process models and, more generally, our view of human rationality (e.g., Gürçay & Baron, 2017; Stanovich & West, 2000). However, in recent years direct experimental testing of the core assumptions has pointed to fundamental issues. Pace the 'bias blind spot' hypothesis, a range of studies have established that often biased reasoners do show bias sensitivity (e.g., Bonner & Newell, 2010; De Neys & Glumicic, 2008; Gangemi, Bourgeois-Gironde, & Mancini, 2015; Pennycook, Trippas, Handley, & Thompson, 2014; Stupple, Ball, Evans, & Kamal-Smith, 2011; but see*

also Aczel, Szollosi, & Bago, 2016; Mata, Ferreira, Voss, & Kollei, 2017; Travers, Rolison, & Feeney, 2016). In these studies, participants are presented with both traditional reasoning problems in which a cued

Bago & De Neys 2019

[188] *"Executive functions (EFs) consist of a family of three, interrelated core skills (inhibitory control, working memory, and cognitive flexibility; Miyake et al., 2000, Diamond, 2013). From those, higher-order EFs are built such as reasoning, problem-solving, and planning (Collins and Koechlin, 2012, Lunt et al., 2012). Inhibitory control involves resisting one's initial impulse or a strong pull to do one thing, and instead act more wisely. Without inhibitory control we would be at the mercy of external stimuli, internal impulses, and habits of thought or action that pull us this way or that. Inhibitory control thus makes it possible for us to choose how we react and to change how we behave rather than being 'unthinking' creatures of habit or impulse (Diamond, 2013). It is critical for avoiding social faux pas and for a civil society where people abide by rules and norms. It is difficult to think of any aspect of life where having the presence of mind to wait before speaking or acting, giving a considered response rather than an impulsive one, being able to stay focused despite distraction, and resisting temptations to do inappropriate, ill-advised, self-destructive or illegal things would not be beneficial. Working memory (WM) involves more than holding information in mind. It involves doing that while performing one or more mental operations. It is needed, for example, for re-ordering the items you are holding in mind or seeing how they relate to one another ('working with' the information you are holding in mind; Baddeley and Hitch, 1994, Smith and Jonides, 1999) and also for remembering your question or comment while following an ongoing discussion or for holding in mind what you were about to do when something arises that must be dealt with first (D'Esposito and Postle, 2015). WM is critical for reasoning and problem-solving for they require holding lots of information in mind, exploring their interrelations, and then perhaps dis-assembling those combinations and re-combining the elements in new ways. WM is necessary for making sense of anything that unfolds over time for that always requires holding in mind what happened earlier and relating that to what is happening now (e.g., following a lecture or conversation, relating what you are reading now to what you read earlier, or understanding the relation between a later effect and an earlier cause).*

Cognitive flexibility refers to the ability to flexibly adjust to changed demands or priorities, to look at the same thing in different ways or from different perspectives (as required for set shifting or task switching; Allport and Wylie, 2000, Kiesel et al., 2010, Monsell, 2003, Vandierendonck et al., 2010). If one way of solving a problem isn't working, one needs cognitive flexibility to 'think outside the box,' that is, to find other ways of conceiving of the problem or of attacking it. Such flexibility is needed for meeting novel, unanticipated challenges and for seizing opportunities when they unexpectedly arise.

EFs are predictive of achievement, health, wealth, and quality of life throughout life, often more so than IQ or socioeconomic status (SES; e.g., Moffitt et al., 2011, Moffitt, 2012). They are more critical for school readiness than IQ or entry-level reading or math (Alloway et al., 2005, Blair, 2002, Blair and Razza, 2007, Carlson and Moses, 2001, Hughes and Ensor, 2008, Morrison et al., 2010). They are predictive of success throughout the school years from preschool through university (often more so than IQ [Duckworth and Seligman, 2005, Alloway and Alloway, 2010, Borella et al., 2010, Duncan et al., 2007, Fiebach et al., 2007, Gathercole et al., 2004, Loosli et al., 2012, McClelland et al., 2007, Nicholson, 2007, Savage et al., 2006])."

Diamond & Ling 2016

[189] "*Recent studies reported that training of working memory may improve performance in the trained function and beyond. Other executive functions, however, have been rarely or not yet systematically examined. The aim of this study was to test the effectiveness of inhibitory control (IC) training to produce true training-related function improvements in a sample of 122 healthy adults using a randomized, double-blind pretest/posttest/follow-up design. Two groups performed either adaptive (training group) or non-adaptive (active control) versions of go/no-go and stop-signal tasks for 3 weeks. Training gains as well as near-transfer to an untrained Stroop task and far-transfer to psychometric fluid intelligence were explored. Although the adaptive group could substantially improve overall IC task performance after training, no differences to the active control group occurred, neither at posttest nor at follow-up testing. A large decrease in response latency from pre- to posttest (and from pretest to 4 months' follow-up testing) was found when the training group was compared to the passive control group, which, however, does not sufficiently control for possible confounds. Thus, no conclusive evidence was found that this performance increase mirrors a true increase in IC function. The fact that training improvement was mainly related to response latency may indicate that individuals were more focused on performance gains in the prepotent go trials but less on the stop trials to meet the requirements*

of the tasks as well as possible. The challenges for response inhibition training studies are extensively discussed. (PsycINFO Database Record (c) 2019 APA, all rights reserved)."

Enge et al. 2014

[190] "*Objectives: Inhibitory control training has been hypothesised as a technique that will improve an individual's ability to overrule impulsive reactions in order to regulate behaviour consistent with long-term goals. Methods: A meta-analysis of 19 studies of inhibitory control training and health behaviours was conducted to determine the effect of inhibitory control training on reducing harmful behaviours. Theoretically driven moderation analyses were also conducted to determine whether extraneous variables account for heterogeneity in the effect; in order to facilitate the development of effective intervention strategies. Moderators included type of training task, behaviour targeted, measurement of behaviour and training duration. Results: A small but homogeneous effect of training on behaviour was found, $d^+ = 0.378$, $CI_{95} = [0.258, 0.498]$. Moderation analyses revealed that the training paradigm adopted, and measurement type influenced the size of the effect such that larger effects were found for studies that employed go/no-go (GNG) training paradigms rather than stop-signal task paradigms, and objective outcome measures that were administered immediately yielded the largest and most consistent effects on behaviour. Conclusions: Results suggest that GNG inhibitory control training paradigms can influence health behaviour, but perhaps only in the short-term. Future research is required to systematically examine the influence of training duration, and the longevity of the training effect. Determining these factors could provide the basis for cost-effective and efficacious health-promoting interventions.*"

Allom, Mullan & Hagger 2016

[191] "*The 'Executive Functions' (EFs) of inhibitory control, working memory, and cognitive flexibility enable us to think before we act, resist temptations or impulsive reactions, stay focused, reason, problem-solve, flexibly adjust to changed demands or priorities, and see things from new and different perspectives. These skills are critical for success in all life's aspects and are sometimes more predictive than even IQ or socioeconomic status. Understandably, there is great interest in improving EFs. It's now clear they can be improved at any age through training and practice, much as physical exercise hones physical fitness. However, despite claims to the contrary, wide transfer does not seem to occur and 'mindless' aerobic exercise does little to improve EFs. Important questions remain: How much can EFs be improved (are benefits only superficial) and how long can benefits be sustained? What are the best methods for improving EFs? What about an approach accounts for its success? Do the answers to these differ by individual characteristics such as age or gender? Since stress, sadness, loneliness, or poor health impair EFs, and the reverse enhances EFs, we predict that besides directly train EFs, the most successful approaches for improving EFs will also address emotional, social, and physical needs.*"

Diamond & Ling 2016

[192] "*To provide evidence for the role of inhibitory control in overcoming deductive reasoning errors, we contrasted the effect of two types of training on the ability to perform deductive reasoning tasks. In one condition, participants were trained to inhibit the perceptual matching bias. In the other condition, participants received training focusing on explaining the underlying logic of the task. Importantly, participants were trained on a different deductive task (i.e., the Wason task, Wason, 1968) than the one performed pre- and post-training (i.e., the perceptual matching bias task, Evans, 1998). The effects of the two types of training were compared to a test-retest control condition in which participants simply performed the perceptual matching task two times. Participants who were trained to inhibit the perceptual matching heuristic were the only ones who succeeded to overcome their deductive reasoning errors. This finding suggests that logical reasoning errors are not due to a lack of logic (or experience) but to a default to inhibit a misleading heuristic. In a follow-up PET (positron emission tomography) imaging study in which we compared the cerebral activation before and after the participants were trained in inhibiting the perceptual matching bias, we observed that the brain activation shifted from the posterior perceptual regions pre-training to prefrontal executive regions post-training. This is the first micro-longitudinal neuroimaging study of deductive reasoning and it provides the first evidence that inhibitory control was critical to reason logically. Note that this brain imaging study on reasoning errors correction was conducted on a sample of only eight participants but the strength of these results stem from the fact that the participants were their own controls in the pre-post training comparison."*

Houdé & Borst 2015

---

193 "*Training interventions to improve decision-making, to date, have met with limited success mostly in specific domains. Training can be very effective when accuracy requires experts to recognize patterns and select an appropriate response, such as in weather forecasting, firefighting, and chess (Phillips, Klein, & Sieck, 2004). By contrast, even highly trained professionals are less accurate than very simple mathematical models in other domains such as parole decisions, personnel evaluations, and clinical psychological testing (Dawes, Faust, & Meehl, 1989). Whether domain- specific expertise is achievable appears to be contingent on external factors such as the prevalence of clear feedback, the frequency of the outcome being judged, and the number and nature of variables that determine that outcome (Harvey, 2011; Kohler, Brenner, & Griffin, 2002).*

Evidence that training effectively improves general decision-making ability is inconclusive at present (Arkes, 1991; Milkman, Chugh, & Bazerman, 2009; Phillips et al., 2004). Weather forecasters are well calibrated when predicting the chance of precipitation (Murphy & Winkler, 1974), for example, but are overconfident in their answers to general knowledge questions (Wagenaar & Keren, 1986). Even within their domain of expertise, experts struggle to apply their training to new problems. Philosophers trained in logic exhibit the same preference reversals in similar moral dilemmas as academics without logic training (Schwitzgebel & Cushman, 2012), and physicians exhibit the same preference reversals as untrained patients for equivalent medical treatments when those treatments are framed in terms of survival or mortality rates (McNeil, Pauker, Sox, & Tversky, 1982). Several studies have shown that people do not apply their training to unfamiliar and dissimilar domains because they lack the necessary metacognitive strategies to recognize underlying problem structure (for reviews, see Barnett & Ceci, 2002; Reeves & Weisberg, 1994; Willingham, 2008)."

Merewedge et al. 2015

194 "*From what we have learned about decision debiasing it seems that it is not enough to educate people about the existence of biases and their functionality; they also need to acquire specific debiasing strategies to cope with these challenges. What makes debiasing even more difficult is that decision makers have to recognize the situations in which they need to use the strategies they learned. This task requires transfer from the rule they learned during the training event to the test situation or (preferably) to any analogous real-life situation. The question is how to train people on an abstract rule that they would apply in various relevant situations. In studies of reasoning, some evidence indicates that practicing only abstract rules can improve performance on specific problems. For example, Fong et al. (1986) found that after training on the law of large numbers where the participants were taught about statistical notions such as sample, population and variability, they were better at reasoning about various uncertainty-related problems, such as slot machines, lotteries, or athletic performance. Similarly, undergraduate and graduate training in psychology and social sciences (Lehman et al., 1988; Lehman and Nisbett, 1990) has been found to increase the students' ability in reasoning about everyday problems involving uncertainty. Nevertheless, it is hard to assume that these people relied only on their abstract knowledge for the new cases and did not benefit from the concrete examples (e.g., the urn problem demonstrations in the study of Fong et al., 1986) use during the training. Closer examination of these results suggests that it is easier to apply the abstract rules in cases with matching superficial features. For example, Cheng et al. (1986) showed that abstract training of the obligation rule ('If precondition P is satisfied, action A must be taken') improves performance on Wason's (1966) four-card problem, but only on those versions of the task where the obligation rule could be used in the task. Fong and Nisbett (1991) taught their participants about the law of large numbers in either one of two domains and they were tested on both domains. Although immediately after the training, they found no effect of domain, 2 weeks later the participants could perform better in the domain they were taught in. In their summary, Smith et al. (1992) suggest that in situations where more than one mechanism is involved, reasoning might rely on hybrid instance-rule mechanisms. Therefore, superficial similarity between the learned instance and the target case can facilitate rule-application. This suggestion is in accord with studies of problem solving where it is assumed that a major cause of failures to transfer the relevant rule to analogous situations is the greater attention people pay to the salient and superficial details at the time of learning and that they will apply the learned principles in the test situation to the degree that it shares those contextual features (Holyoak and Koh, 1987; Ross, 1987).*"

Aczel et al. 2015

195 *"In each case, expertise in a domain helps people develop a sensitivity to patterns of meaningful information that are not available to novices… For example, electronics technicians were able to reproduce large portions of complex circuit diagrams after only a few seconds of viewing ; novices could not. The expert circuit technicians chunked several individual circuit elements (eg. resistors and capacitors) that performed the*

function of an amplifier. By remembering the structure and function of a typical amplifier, experts were able to recall the arrangement of the many individual circuit elements comprising the '« amplifier chunk' ». ... physicists recognize problems of river currents and problems of headwinds and tailwinds in airplanes as involving similar mathematical principles, such as relative velocities." »

(Bransford et al 2000)

196 *"Experts notice features and meaningful patterns of information that are not noticed by novices.*
Experts have acquired a great deal of content knowledge that is organized in ways that reflect a deep understanding of their subject matter.
Expert knowledge cannot be reduced to sets of isolated facts or propositions but, instead, reflects contexts of applicability ; that is, the knowledge is '« conditionalized' » on a set of circumstances.
Experts are able to flexibly retrieve important aspects of their knowledge with little attentional effort.
Though experts know their disciplines thoroughly, this does not guarantee that they are able to teach others.
Experts have varying levels of flexibility in their approach to new situations."

(Bransford et al 2000)

197 "*...when the conceptual knowledge of subjects is controlled (sometimes called the level of expertise) in relation to specific fields (football, chess, dinosaurs), it becomes impossible to highlight a development: those who know more do better than others, regardless of age (Chi and Ceci 1987; Yates & Chandler 1991).* »

Fayol & Monteil 1994

198 *"A strategy is an integrated sequence, more or less long and complex, of procedures selected for a goal in order to achieve optimal performance. It may involve very general procedures - for example, the idea that it is necessary to plan intentionally to achieve a goal - or very specific procedures - for example, asking oneself questions to ensure that one has understood a text (Nisbet & Shucksmith, 1986).* »

Fayol & Monteil 1994

199 *"The most influential factor in mobilizing procedures seems to be prior knowledge of the subject matter in relation to the conceptual area being discussed*. »

Fayol & Monteil 1994

200 "*Studies of the Philosophy for Children program may be taken as typical. Two researchers identified eight studies that evaluated academic outcomes and met minimal research-design criteria. (Of these eight, only one had been subjected to peer review.) Still, they concluded that three of the eight had identifiable problems that clouded the researchers' conclusions. Among the remaining five studies, three measured reading ability, and one of these reported a significant gain. Three studies measured reasoning ability, and two reported significant gains. And, two studies took more impressionistic measures of student's participation in class (e.g., generating ideas, providing reasons), and both reported a positive effect*."

Willingham, 2007

201 "*Despite the difficulties and general lack of rigor in evaluation, most researchers reviewing the literature conclude that some critical thinking programs do have some positive effect. But these reviewers offer two important caveats. First, as with almost any educational endeavor, the success of the program depends on the skill of the teacher. Second, thinking programs look good when the outcome measure is quite similar to the material in the program. As one tests for transfer to more and more dissimilar material, the apparent effectiveness of the program rapidly drop*s."

Willingham, 2007

202 "*If you remind a student to 'look at an issue from multiple perspectives' often enough, he will learn that he ought to do so, but if he doesn't know much about an issue, he can't think about it from multiple perspectives. You can teach students maxims about how they ought to think, but without background knowledge and practice, they probably will not be able to implement the advice they memorize*."

Willingham, 2007

---

[203] "*Can critical thinking actually be taught? Decades of cognitive research point to a disappointing answer: not really. People who have sought to teach critical thinking have assumed that it is a skill, like riding a bicycle, and that, like other skills, once you learn it you can apply it in any situation. Research from cognitive science shows that thinking is not that sort of skill. The processes of thinking are intertwined with the content of thought (that is, domain knowledge)."*

Willingham 2007

[204] *"If knowledge of how to solve a problem never transferred to problems with new surface structures, schooling would be inefficient or even futile—but of course, such transfer does occur. When and why is complex, but two factors are especially relevant for educators: familiarity with a problem's deep structure and the knowledge that one should look for a deep structure*."

Willingham, 2007

[205] *"Here's an example: A treasure hunter is going to explore a cave up on a hill near a beach. He suspected there might be many paths inside the cave so he was afraid he might get lost. Obviously, he did not have a map of the cave; all he had with him were some common items such as a flashlight and a bag. What could he do to make sure he did not get lost trying to get back out of the cave later? The solution is to carry some sand with you in the bag, and leave a trail as you go, so you can trace your path back when you're ready to leave the cave. About 75 percent of American college students thought of this solution—but only 25 percent of Chinese students solved it.6 The experimenters suggested that Americans solved it because most grew up hearing the story of Hansel and Gretel, which includes the idea of leaving a trail as you travel to an unknown place in order to find your way back. The experimenters also gave subjects another puzzle based on a common Chinese folk tale, and the percentage of solvers from each culture reversed. (To read the puzzle based on the Chinese folk tale, and the tale itself, go to www.aft.org/pubs-reports/american_educator/index.htm.) It takes a good deal of practice with a problem type before students know it well enough to immediately recognize its deep structure, irrespective of the surface structure, as Americans did for the Hansel and Gretel problem*."

Willingham, 2007

[206] "*They are little chunks of knowledge—like 'look for a problem's deep structure' or 'consider both sides of an issue'—that students can learn and then use to steer their thoughts in more productive directions." Le problème de ce genre de stratégie métacognitive est qu'elle ne peut pas porter beaucoup plus loin qu'à amener à se répéter au bon moment la stratégie métacognitive. 'Thus, a student who has been encouraged many times to see both sides of an issue, for example, is probably more likely to spontaneously think "I should look at both sides of this issue" when working on a problem. … Unfortunately, metacognitive strategies can only take you so far. Although they suggest what you ought to do, they don't provide the knowledge necessary to implement the strategy."*

Willingham, 2007

[207] "*Understanding and using conditional probabilities is essential to scientific thinking because it is so important in reasoning about what causes what. But people's success in thinking this way depends on the particulars of how the question is presented. Studies show that adults sometimes use conditional probabilities successfully, but fail to do so with many problems that call for it. Even trained scientists are open to pitfalls in reasoning about conditional probabilities (as well as other types of reasoning). Physicians are known to discount or misinterpret new patient data that conflict with a diagnosis they have in mind, and Ph.D.- level scientists are prey to faulty reasoning when faced with a problem embedded in an unfamiliar context.*"

Willingham, 2007

[208] "*But critical thinking is very different. As we saw in the discussion of conditional probabilities, people can engage in some types of critical thinking without training, but even with extensive training, they will sometimes fail to think critically. This understanding that critical thinking is not a skill is vital*."

Willingham, 2007

[209] "*It tells us that teaching students to think critically probably lies in small part in showing them new ways of thinking, and in large part in enabling them to deploy the right type of thinking at the right time*."

"*What do all these studies boil down to? First, critical thinking (as well as scientific thinking and other domain-based thinking) is not a skill. There is not a set of critical thinking skills that can be acquired and deployed regardless of context. Second, there are metacognitive strategies that, once learned, make critical thinking more likely. Third, the ability to think critically (to actually do what the metacognitive strategies call for) depends on domain knowledge and practice. For teachers, the situation is not hopeless, but no one should underestimate the difficulty of teaching students to think critically*."

Willingham, 2007

[210] "*Numerous qualitatively different forms of outcome evaluations for thinking courses provide substantial evidence for the conclusion that it is possible to use education to improve the ability to think critically, especially when instruction is specifically designed to encourage transfer of these skills to different situations and different domains of knowledge*."

Halpern 2013

[211] "*Critical thinking does not automatically result as a byproduct of standard instruction in a content area. Critical thinking instruction needs to focus overtly and self-consciously on the improvement of thinking, and the learning experience needs to include multiple examples across domains in order to maximize transfer*."

Halpern 2013

[212] "*The best way to promote the kind of transfer I am advocating is with the conscious and deliberate use of the skills that are learned in a wide variety of contexts."*

Halpern 2013

[213] "*Ideally critical thinking skills should be used to recognize and resist unrealistic campaign promises, circular reasoning, faulty probability estimates, weak arguments by analogy, or language designed to mislead whenever and wherever it is encountered. Critical thinkers should be able to solve or offer reasonable solutions to real world problems, whether it is the problem of nuclear war or how to set up a new computer. These skills should also be long lasting and used for the many decades of critical thinking that most of us will face…*"

Halpern 2013

[214] *"Our working definition for the purposes of this review is that thinking skills interventions are approaches or programmes which identify for learners translatable, mental processes and/or which require learners to plan, describe and evaluate their thinking and learning. These can therefore be characterised as approaches or programmes which:*

- *require learners to articulate and evaluate specific learning approaches; and/or*
- *identify specific cognitive, and related affective or conative processes that are amenable to instruction."*

Higgins et al. 2005

[215] "*Studies were selected for the meta-analysis if they had sufficient quantitative data to calculate an effect size (relative to a control or comparison group of pupils) and if the number of research subjects was greater than 10. Effect sizes were calculated from the reported data and combined statistically using quantitative synthesis*." Le résultat est que "*Twenty-nine studies were identified which contained quantitative data on pupils' attainment and attitudes suitable for meta-analysis. The studies come from a range of countries around the world with half set in the US and UK. The studies broadly cover the ages of compulsory schooling (5–16) and include studies set in both primary and secondary schools. A number of named thinking skills interventions are included, such as Feuerstein's instrumental enrichment (FIE) and cognitive acceleration through science education (CASE) as well as studies which report a more general thinking skills approach (such as the development of metacognitive strategies)."*

Higgins et al. 2005

[216] *"Twentieth-century psychologists have been pessimistic context of solving recurrent everyday problems. These rule about teaching reasoning, prevailing opinion suggesting that people may possess only domain-specific rules, rather than abstract rules; this would mean that training a rule in one domain would not produce*

generalization to other domains. … We propose an alternative view that is close to the pre-20th- century one: people do make use of inferential rules and these can be readily taught. In fact, rules that are extensions of naturally induced ones can be taught by quite abstract means. This description does not apply to formal, deductive logical rules or to most other purely syntactic rule systems, however. Instead, the types of inferential rules that people use naturally and can be taught most easily are a family of pragmatic inferential rule systems that people induce in the context of solving recurrent everyday problems (4). These rule systems are abstract in as much as they can be used in a wide variety of content domains, but their use is confined to certain types of problem goals and particular types of relations between events. They include 'causal schemas' (5), 'contractual schemas,' such as the rules underlying permission and obligation in the social sphere, and 'statistical heuristics,' used in the evaluation of evidence, such as qualitative, intuitive versions of the law of large numbers."

Nisbett et al. 1986

217 "*The data base students acquire in school ought to inform their thinking in other subjects and in life outside the schoo*l."

Perkins & Salomon 1988

218 "*Analogy and transfer of learning … are both concerned with the same basic questions: how can what has been learnt in a certain situation be generalized to new situations and what conditions are necessary and/or sufficient to generate transfer from one situation to anothe*r?"

Gamo, Sander & Richard 2010

219 "*Since transfer between tasks is a function of the similarity by transfer trasks and learning experiences, an important strategy for enhancing transfer from schools to other settings may be to better understand the non-school environments in which students must function."*

Bransford et al. 2000

220 "*The transfer literature suggests that the most effective transfer may come from a balance of specific examples and general principles, not from either one alone*."

Bransford et al. 2000

221 "*Instead of worrying about which is more important – local knowledge or the more general, transferable aspects of knowledge – we should recognize the synergy of local and more general knowledge. … students who lack an understanding of key mathematical concepts will not gain much from the general strategy of trying to define and represent a problem well before they start. But students who lack the habit of trying to define and represent a problem well will often misuse the mathematical concepts they know when the problem is not routine. … Proper attention to transfer will make the best of both for the sake of deeper and broader skill, knowledge and understanding*."

Salomon & Perkins 1987

222 "*Humans are not naturally critical. Indeed, like ballet, critical thinking is a highly contrieved activity. Running is natural; night club dancing is less so; but ballet is something people can only do well with many years of painful, expensive, dedicated training.*"

Van Gelder 1995

223 "*Thus critical thinking cannot be treated just as a kind of gloss on educational content made up to other 'real' subjects. Students will not become excellent critical thinkers merely by studying history, marketing or nursing, even if the instruction is given a 'critical' emphasis (as it should be). Critical thinking must be studies and practiced in its own right; it must be an explicit part of the curriculum.*"

Van Gelder 1995

224 "*Critical thinking skills are, by definition, ones that apply in a wide range of domains, contexts, and so on, and there is plenty of territories in which they can fail to transfer*."

Van Gelder 1995

---

[225] "*There is little evidence of the benefit of teaching metacognitive approaches in 'learning to learn' or 'thinking skills' sessions. Pupils find it hard to transfer these generic tips to specific tasks. Self-regulated learning and metacognition have often been found to be context-dependent, so how you best plan in Key Stage 2 art may have significant differences to planning strategies in GCSE maths. This means that pupil who shows strong self-regulated learning and metacognitive competence in one task or subject domain may be weak in another, and metacognitive strategies may or may not be effective, depending on the specific task, subject, or problem tackled. This does not, however, mean that metacognitive knowledge and skills will automatically develop through content knowledge teaching. That being said, over time, metacognition can become more generic, and older metacognitive learners can possess an array of strategies that they then judiciously apply across a range of contexts and to a range of tasks. This maturation also includes the development of a growing understanding of when to use what strategies, or when good strategies may be missing in the learner's repertoire*."

EEF, https ://educationendowmentfoundation.org.uk/tools/guidance-reports/metacognition-and-self-regulated-learning

[226] "*In contrast to approaches that emphasize explicit instruction as the most powerful tool in developing critical thinking and writing (Larson, 100 Britt, & Kurby, 2009; Marin & Halpern, 2011; Schworm & Renkl, 2007), our approach is experiential in its pedagogical emphasis and microgenetic with regard to research methodology (Kuhn, 1995). By observing students engaging in guided practice over a period of time, we believe we can learn something about what develops and how*."

Kuhn, Hemberger & Khait 2015

[227] "*Our earlier work (Felton & Kuhn, 2001; Kuhn & Udell, 2003; Kuhn, Goh, Iordanou, & Shaenfield, 2008) documented that young adolescents engaged in argumentation typically concentrate their attention on exposition of their own claims, essentially ignoring the opponent's position. Thus, the initial goals of our programme are to encourage attention to the other's position and to enhance ability and disposition to address it with the goal of weakening it, or in other words, to engage in counter argumentation. Once this goal is achieved, our emphasis shifts to the use of evidence to strengthen and weaken claims. By securing answers to their own self-generated questions, students begin to contribute to the set of evidence bearing on the topic that we provide initially, and this evidence comes to play an increasing role in their argumentation.*"

Kuhn, Hemberger & Khait 2015

[228] "*In their discourse with one another, students will not be able to generate rich arguments and counterarguments in a vacuum. They need to bring to bear information relevant to the topic to inform their reasoning. Depending on the topic, they will likely already have some degree of related knowledge that they are ready to summon to support their arguments (or to weaken those of their opponents). But they will need more knowledge than they have at the outset. One approach we might take is to ask students to begin their work on a topic by reading material about it that they can then draw on in their argumentation. A problem with this approach is that students don't yet appreciate the purpose that this information serves. In a word, it provides answers to questions they don't yet have. As a result, they fail to see its point. They are thus likely to approach such reading disinterestedly, as just another reading assignment to be completed and most likely forgotten. A small dose of initial reading can be productive, to heighten initial interest in the topic, but at the outset we employ it sparingly out of concern that a deluge of information up front not only is met with disinterest but can shut down students' own thinking and inquisitiveness about a topic. Therefore, we let students' own ideas dominate at the beginning of their engagement with a new topic, encouraging them to articulate and share with one another their ideas about the topic. And, we have found, they do have lots of ideas to share, even in the case of topics outside the range of their immediate experience that we might expect them to know little about. Still, students' discourse stands to be informed and enriched by information bearing on the topic that likely will be new to them. An effective way to do this, we have found, is to create a need for the information they acquire. Rather than provide answers to questions students don't have, we let them first formulate the questions. In this way, we allow students to first see how such information could be useful in achieving their discourse objectives, and then we assist them in securing it. Hence, after introducing a few basic questions and answers regarding the topic, we invite students to generate questions of their own, the answers to which they think might be helpful to them. By the next session, we then make available brief factual answers to these questions (which*

sometimes students assist in obtaining), and the resulting question-and-answer 'evidence set' (containing all questions and answers) we assemble for them as a written document that is made available for use by the entire class throughout their work on the topic. The point is for students not just to acquire information but to see its value and therefore be disposed to apply it. With practice, we have found, students do in fact in time make much use of this information, coming to recognize it as playing a critical role in their discourse."

Kuhn, Hemberger & Khait 2015

229 "*According to the argumentative theory of reasoning, the main function of human reasoning is to exchange arguments with others (Mercier & Sperber, 2011). This theory predicts that reasoning should yield better outcomes when it is used in conversation than in solitary ratiocination. It is thus well supported by the comparison of individual and group performance on intellective tasks (as well as on a variety of others tasks, see Mercier, 2016; Mercier & Sperber, 2011)*."

Boku, Yama & Mercier 2018

230 *"Finally, participants who change their mind following group discussion tend to perform better on transfer problems than they had before the discussion (Laughlin, 2011; Trouche et al., 2014). Although the argumentative theory of reasoning does not predict that such transfer will always take place, when it does it is a sign that participants have understood why they should change their mind to accept the correct answer."*

Boku, Yama, Mercier 2018

231 "*The most straightforward consequence of the combination of myside bias and laziness that characterizes argument production is that when people reason on their own, reasoning typically fails to correct misguided intuitions (e.g., [24]). For instance, most participants who tackle the bat and ball on their own fail to provide the right answer, although it is mathematically trivial [25]. Solitary reasoning can even lead to an accumulation of arguments supporting our opinions, most of them poorly examined, leading to overconfidence [22, 23] and polarization [26]. Solitary reasoning also provides people with excuses to engage in morally dubious behavior [27]."*

Mercier 2016

232 "*The argumentative theory of reasoning predicts that people evaluate others' arguments in a way that is objective and demanding. However, in contrast with argument production, which has immediately perceivable effects (the arguments produced), argument evaluation is more difficult to measure. The evaluation of the strength of an argument is only one of the factors that determine whether its conclusion is accepted. Prior beliefs regarding the plausibility of the argument's conclusion and the trustworthiness of the argument's source also have to be taken into account. As a result, an argument's conclusion might be rejected, not because the argument was not evaluated properly, but because it failed to generate enough conflict with one's prior beliefs to lead to a change of mind*."

Mercier 2016

233 "*The study of persuasion and attitude change has shown that when participants are given arguments on issues that they care about, good arguments are much more effective at changing their minds than weak ones [12]. When participants evaluate everyday arguments, they react appropriately to variations in argument strength, whether strength is measured by Bayesian modeling [49–52], norms of classical logic [53], or norms of argumentation fallacies [54–57]. For instance, arguments from authority are deemed potentially fallacious if they come from non-expert sources or from sources with conflicts of interest, and participants appropriately discount such arguments [55]."*

Mercier 2016

References

1. Introduction

- Acerbi, A. (2019). Cognitive attraction and online misinformation. *Palgrave Communications*, 5(1), 15.
- Evans, J. S. B. T. (2012). Dual process theories of deductive reasoning: facts and fallacies. *The Oxford handbook of thinking and reasoning*, 115-133.
- Gelder, T. V. (2005). Teaching critical thinking: Some lessons from cognitive science. *College teaching*, 53(1), 41-4
- Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge university press.
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American psychologist*, *53*(4), 449.
- Halpern, D. F. (2013). *Thought and knowledge: An introduction to critical thinking*. Psychology Press.
- Howells, K. (2018). *The future of education and skills: education 2030: the future we want*.
- Kahneman, D. (2002). Maps of bounded rationality: A perspective on intuitive judgment and choice. *Nobel prize lecture*, *8*, 351-401.
- Kuhn, D. (1999). A developmental model of critical thinking. *Educational researcher*, *28*(2), 16-46.
- Nisbett, R. E. (2015). *Mindware: Tools for smart thinking*. Farrar, Straus and Giroux.
- Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28(3), 369-381.
- Shafir, E., & LeBoeuf, R. A. (2002). Rationality. *Annual review of psychology*, *53*(1), 491-517.
- Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning*. Psychology Press.
- Stanovich, K. E. (2009). *What intelligence tests miss: The psychology of rational thought*. Yale University Press.
- Stanovich, K. E., & Stanovich, P. J. (2010). *A framework for critical thinking, rational thinking, and intelligence. Innovations in educational psychology: Perspectives on learning, teaching and human development*, 195-237.
- Van Gelder, T. (2005). Teaching critical thinking: Some lessons from cognitive science. *College teaching*, 53(1), 41-48.
- Whitworth, A. (2009). *Information obesity*. Elsevier.

2. Towards an operational definition of CT

Cited references

- Bailin, S. (2002). Critical thinking and science education. *Science & Education*, 11(4), 361–375.
- Bailin, S., & Battersby, M. (2016). *Reason in the balance: An inquiry approach to critical thinking*. Hackett Publishing.
- Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999). Conceptualizing critical thinking. *Journal of curriculum studies*, *31*(3), 285-302.
- Beran, M. J., Perner, J., & Proust, J. (Eds.). (2012). *Foundations of metacognition*. Oxford University Press.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain.* New York: David McKay Co Inc.
- Boisvert, J. (2016). La formation de la pensée critique: deuxième partie. Bulletin de La Documentation Collégiale N° 15, Janvier 2016.
- Bronner, G. (2013). *La démocratie des crédules*. Presses universitaires de France.
- Cosmides, L., & Tooby, J. (1994). Better than rational : Evolutionary psychology and the invisible hand. *The American Economic Review,84*(2), 327-332.
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *cognition*, *58*(1), 1-73.
- Dewey, J. (1933). *How we think*. New York: Health and Company.
- Dewey, J. (1997). Experience and education. 1938. *New York: First Touchstone Edition*, 64-67.
- Elder, L., & Paul, R. (2010). Critical Thinking: Competency Standards Essential for the Cultivation of Intellectual Skills, Part 1. *Journal of Developmental Education*, *34*(2), 38-39.
- Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. Educational Researcher, 18(3), 4–10
- Ennis, R. H. (1991). Critical thinking: A streamlined conception. *Teaching philosophy*, *14*(1), 5-24.
- Ennis, R. H. (2011). The nature of critical thinking: An outline of critical thinking dispositions and abilities. *University of Illinois*, 2-4.
- Ennis, R. H. (2016). Definition: A Three-Dimensional Analysis with Bearing on Key Concepts. In Proceedings of the Ontario Society for the Study of Argumentation Conference (Vol. 11, p. 150).
- Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science*, *8*(3), 223-241.
- Facione, P. A. (1990). Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. Research Findings and Recommendations.
- Facione, P. A. (2011). Critical thinking: What it is and why it counts. *Insight assessment*, *2007*(1), 1-23.
- Foundation for critical thinking, https://www.criticalthinking.org/pages/defining-critical-thinking/766

- Gelder, T. V. (2005). Teaching critical thinking: Some lessons from cognitive science. College teaching, 53(1), 41-48.
- Gentner, D. (1989). Analogical learning. *Similarity and analogical reasoning*, *199*.
- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". *European review of social psychology*, *2*(1), 83-115.
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky.
- Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. *Blackwell handbook of judgment and decision-making*, *62*, 88.
- Gigerenzer, G., Hoffrage, U., & Goldstein, D. G. (2008). *Fast and frugal heuristics are plausible models of cognition: Reply to Dougherty*, Franco-Watkins, and Thomas (2008).
- Gigerenzer, G., & Todd, P. M. (1999). ABC Research Group. *Simple heuristics that make us smart*.
- Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge university press.
- Glaser, E. M. (1941). *An experiment in the development of critical thinking*. Teachers College, Columbia University.
- Gopnik, A., Schulz, L., & Schulz, L. E. (Eds.). (2007). *Causal learning: Psychology, philosophy, and computation*. Oxford University Press.
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American psychologist*, *53*(4), 449.
- Halpern, D. F. (1999). Teaching for critical thinking: Helping college students develop the skills and dispositions of a critical thinker. *New directions for teaching and learning*, *1999*(80), 69-74.
- Halpern, D. F. (2007). The nature and nurture of critical thinking. *Critical thinking in psychology*, 1-14.
- Halpern, D. F. (2013). *Thought and knowledge: An introduction to critical thinking*. Psychology Press.
- Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W. E., & Moore, T. (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition*, *27*(5), 733-763.
- Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and social psychology Review*, *10*(1), 47-66.
- Hitchcock, D. (2018). Critical thinking. Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/fall2018/entries/critical-thinking/
- Kahan, D. M. (2015). The politically motivated reasoning paradigm, part 1: What politically motivated reasoning is and how to measure it. *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource*, 1-16.
- Kahan, D. M., Jenkins-Smith, H., & Braman, D. (2011). Cultural cognition of scientific consensus. *Journal of risk research*, *14*(2), 147-174.

- Kahan, D. M., Braman, D., Cohen, G. L., Gastil, J., & Slovic, P. (2010). Who fears the HPV vaccine, who doesn't, and why? An experimental study of the mechanisms of cultural cognition. *Law and human behavior*, *34*(6), 501-516.
- Kahan, D. M., & Stanovich, K. (2016). Rationality and belief in human evolution. *Annenberg Public Policy Center Working Paper*, (5)
- Kahneman, D., & Tversky, A. (1996). *On the reality of cognitive illusions*.
- Kahneman, D., Slovic, S. P., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge university press.
- Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review*, *62*(2), 155-179.
- Kuhn, D. (1999). A developmental model of critical thinking. *Educational researcher*, *28*(2), 16-46.
- Kuhn, D. (2018). A role for reasoning in a dialogic approach to critical thinking. Topoi, 37(1), 121-128.
- Lai, E. R. (2011). Critical thinking: A literature review. *Pearson's Research Reports*, *6*, 40-41.
- Lewis, A., & Smith, D. (1993). Defining Higher Order Thinking. *Theory into Practice*, 32(3,), 131–137.
- Lipman, M. (2003). Thinking in education. (2nd. ed.). New York: Cambridge University Press.
- Lipman, M.. (1987). Critical thinking: What can it be? *Analytic Teaching*, 8(1).
- Lipman, M. (1988). Critical thinking: What can it be? Resource Publication, Series 1 No. 1.
- McPeck, J. E. (1990). Critical thinking and subject specificity: A reply to Ennis. *Educational researcher*, *19*(4), 10-12.
- Mercier, H. (2017). How gullible are we? A review of the evidence from psychology and social science. *Review of General Psychology*, *21*(2), 103-122.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory.
- Mercier, H., & Sperber, D. (2017). *The enigma of reason*. Harvard University Press.
- Nisbett, R. E. (2015). *Mindware: Tools for smart thinking*. Farrar, Straus and Giroux.
- Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science*, *238*(4827), 625-631.
- Oreskes, N. (2019). *Why Trust Science?* (Vol. 1). Princeton University Press.
- Origgi, G. (2019). *Reputation: What it is and why it Matters*. Princeton University Press.
- Paul, R. & Elder. L. (2009). *The Miniature Guide to Critical Thinking Concepts and Tools*, Foundation for Critical Thinking, Dillon Beach, CA, 5th Ed.
- Paul, R. W., Elder, L., & Bartell, T. (1997). *California teacher preparation for instruction in critical thinking: Research findings and policy recommendations*.
- Siegel, H. (2010). Critical thinking. In *International encyclopedia of education* (pp. 141-145). Elsevier Ltd.
- Scriven, M., & Paul, R. (1987). Critical thinking as defined by the National Council for Excellence in Critical Thinking. In *8th Annual International Conference on Critical Thinking and Education Reform, Rohnert Park, CA* (pp. 25-30).
- Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental science, 10*(1), 89-96.

- Stanovich, K. E., & Stanovich, P. J. (2010). A framework for critical thinking, rational thinking, and intelligence. *Innovations in educational psychology: Perspectives on learning, teaching and human development*, 195-237.
- Sternberg, R. J. (1986). Critical Thinking: Its Nature, Measurement, and Improvement.
- Trouche, E., Johansson, P., Hall, L., & Mercier, H. (2016). The selective laziness of reasoning. *Cognitive Science*, *40*(8), 2122-2136.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124-1131.
- West, R. F., Toplak, M. E., & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology*, *100*(4), 930.
- Willingham, D. T. (2007). Critical thinking: Why it is so hard to teach? *American federation of teachers summer 2007, pp. 8-19*.
- Willingham, D. T. (2019). How to teach critical thinking. Education: Future Frontiers Occasional Paper Series. Retrieved from Analysis and Policy Observatory Website: https://apo.org.au/node/244676

Additional readings, used for the Report

Sharon Bailin

- Bailin, S. (1987). Critical and creative thinking. *Informal logic*, 9(1).
- Bailin, S., Battersby, M., & Clauss, P. (2011). *Reason in the balance: Teaching critical thinking as dialectical*. 13.
- Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999a). Common misconceptions of critical thinking. *Journal of Curriculum Studies*, 31(3), 269–283.
- Bailin, S., & Siegel, H. (2003). Critical thinking. The Blackwell Guide to the *Philosophy of Education*, 181, 193.
- Bailin, S. (1998). *Skills, generalizibility and critical thinking. In twentieth world congress on philosophy*. Boston: The Paideia Archive.
- Battersby, M., & Bailin, S. (2011). Critical Inquiry: Considering the Context. *Argumentation*, 25(2), 243–253.

Robert Ennis

- Ennis, R. H. (1962). A concept of critical thinking. *Harvard educational review*.
- Ennis, R. H. (1964). A definition of critical thinking. *The Reading Teacher*, *17*(8), 599-612.
- Ennis, R. H. (1993). Critical thinking assessment. *Theory into practice*, *32*(3), 179-186.
- Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational researcher*, *18*(3), 4-10.
- Ennis, R. H. (1984). Problems in testing informal logic critical thinking reasoning ability. *Informal Logic*, *6*(1).
- Ennis, R. H. (1985). A logical basis for measuring critical thinking skills. *Educational leadership*, *43*(2), 44-48.

- Ennis, R. (1991). Critical thinking: A streamlined conception. *Teaching philosophy*, *14*(1), 5-24.
- Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities.
- Ennis, R. H. (1996). Critical thinking dispositions: Their nature and assessability. *Informal logic*, *18*(2).
- Ennis, R. H. (1987). Critical thinking and the curriculum. *Thinking skills instruction: Concepts and techniques*, 40-48.
- Kennedy, M., Fisher, M. B., & Ennis, R. H. (1991). Critical thinking: Literature review and needed research. *Educational values and cognitive instruction: Implications for reform*, *2*, 11-40.
- Norris, S. P., & Ennis, R. H. (1989). *Evaluating Critical Thinking. The Practitioners' Guide to Teaching Thinking Series*. Critical Thinking Press and Software, Box 448, Pacific Grove, CA 93950-0448; tele.
- Ennis, R. H. (1990). The extent to which critical thinking is subject-specific: Further clarification. *Educational researcher*, *19*(4), 13-16.
- Ennis, R. (2011). Critical thinking: Reflection and perspective Part II. *Inquiry: Critical thinking across the Disciplines*, *26*(2), 5-19.
- Ennis, R. H. (1998). Is critical thinking culturally biased? *Teaching philosophy*, *21*(1), 15-33.
- Ennis, R. H. (2015). Critical thinking: A streamlined conception. In *The Palgrave handbook of critical thinking in higher education* (pp. 31-47). Palgrave Macmillan, New York.
- Ennis, R. (1997). Incorporating critical thinking in the curriculum: An introduction to some basic issues. *Inquiry: Critical Thinking across the disciplines*, *16*(3), 1-9.

Peter Facione

- Facione, P. A., Sanchez, C. A., Facione, N. C., & Gainen, J. (1995). The disposition toward critical thinking. *The Journal of General Education*, 1-25.
- Facione, N. C., Facione, P. A., & Sanchez, C. A. (1994). Critical thinking disposition as a measure of competent clinical judgment: The development of the California Critical Thinking Disposition Inventory. *Journal of Nursing Education*, *33*(8), 345-350.
- Facione, P. A. (2000). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. *Informal logic*, *20*(1).
- Facione, P. A., & Facione, N. C. (1994). Holistic critical thinking scoring rubric. *Retrieved July*, *19*, 2007.
- Giancarlo, C. A., & Facione, P. A. (2001). A look across four years at the disposition toward critical thinking among undergraduate students. *The Journal of General Education*, *50*(1), 29-55.
- Facione, N. C., & Facione, P. A. (2008). Critical thinking and clinical judgment. *Critical thinking and clinical reasoning in the health sciences: A teaching anthology*, 1-13.
- Facione, P. A., Facione, N. C., & Giancarlo, C. A. (1997). Professional judgment and the disposition toward critical thinking. *Retrieved August*, *1*, 2004.
- Facione, P. A. (1990). The California Critical Thinking Skills Test--College Level. Technical Report# 2. Factors Predictive of CT Skills.

- Facione, P. A. (1990). The California Critical Thinking Skills Test--College Level. Technical Report# 1. Experimental Validation and Content Validity.
- Facione, P. A. (1984). Toward a theory of critical thinking. *Liberal Education, 70*(3), 253-61.
- Facione, P. A. (1991). Using the California Critical Thinking Skills Test in Research, Evaluation, and Assessment.
- Facione, P. A., & Facione, N. C. (2007). Talking critical thinking. *Change: The Magazine of Higher Learning*, *39*(2), 38-45.

Diane Halpern / related to Halpern Critical Thinking Assessment

- Halpern, D. F. (1993). Assessing the effectiveness of critical-thinking instruction. *The Journal of General Education*, 238-254.
- Halpern, D. F. (2014). *Critical thinking across the curriculum: A brief edition of thought & knowledge*. Routledge.
- Halpern, D. F. (1987). Analogies as a critical thinking skill. *Applications of cognitive psychology: Problem solving, education, and computing*, 75-86.
- Halpern, D. F. (2006). Is intelligence critical thinking? Why we need a new definition of intelligence. *Extending intelligence: Enhancement and new constructs, 293310*.
- HAU, K. T., HALPERN, D. F., MARIN-BURKHART, L., HO, T., Ku, Y. L., CHAN, N. M., & LUN, M. C. (2006). Chinese and United States Students' critical thinking: Cross-cultural construct validation of a critical thinking assessment.
- Butler, H. A. (2012). Halpern Critical Thinking Assessment predicts real-world outcomes of critical thinking. *Applied Cognitive Psychology*, *26*(5), 721-729.
- Marin, L. M., & Halpern, D. F. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains. *Thinking Skills and Creativity*, *6*(1), 1-13.
- Butler, H. A., Dwyer, C. P., Hogan, M. J., Franco, A., Rivas, S. F., Saiz, C., & Almeida, L. S. (2012). The Halpern Critical Thinking Assessment and real-world outcomes: Cross-national applications. *Thinking Skills and Creativity*, *7*(2), 112-121.
- Sternberg, R. J., Roediger III, H. L., & Halpern, D. F. (Eds.). (2007). *Critical thinking in psychology*. Cambridge University Press.

Deanna Kuhn

- Dean, D., & Kuhn, D. (2003). Metacognition and Critical Thinking.
- Kuhn, D., & Dean, Jr, D. (2004). Metacognition: A bridge between cognitive psychology and educational practice. *Theory into practice*, *43*(4), 268-273.
- Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science education*, *77*(3), 319-337.
- Kuhn, D. (2015). Thinking together and alone. *Educational researcher*, *44*(1), 46-53.
- Felton, M. K., & Kuhn, D. (2007). "How do I Know?" The Epistemological Roots of Critical Thinking. *Journal of Museum Education*, *32*(2), 101-110.
- Kuhn, D. (1996). Is good thinking scientific thinking. *Modes of thought: Explorations in culture and cognition*, 261-281.
- Kuhn, D. (2018). A role for reasoning in a dialogic approach to critical thinking. *Topoi*, *37*(1), 121-128
- Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological science*, *22*(4), 545-552.

- Kuhn, D. (2005). *Education for thinking*. Harvard University Press.
- Kuhn, D., Hemberger, L., & Khait, V. (2017). Argue with me: Argument as a path to developing students' thinking and writing. Routledge.
- Kuhn, D. (2000). Metacognitive development. Current directions in psychological science, 9(5), 178-181.
- Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. Cognitive development, 15(3), 309-328.
- Kuhn, D. (1991). Education for thinking: What can psychology contribute? Promoting cognitive growth over the life span.
- Kuhn, D. (1992). Thinking as argument. Harvard Educational Review, 62(2), 155-179.
- Kuhn, D., & Udell, W. (2003). The development of argument skills. Child development, 74(5), 1245-1260.

Matthew Lipman

- Lipman, M. (2003). *Thinking in education*. Cambridge University Press.
- Lipman, M. (1974). *Harry Stottlemeier's discovery*.
- Lipman, M. (1985). Thinking skills fostered by philosophy for children. *Thinking and Learning Skill Vol. 1: Relating Instruction to Research*, 83-108.
- Lipman, M. (1984). The Cultivation of Reasoning through Philosophy. *Educational leadership*, *42*(1), 51-56.
- Lipman, M. (1998). Teaching students to think reasonably: Some findings of the Philosophy for Children program. *The Clearing House*, *71*(5), 277-280.
- Lipman, M. (1987). *Some thoughts on the foundations of reflective education*.
- Lipman, M. (1993). Promoting better classroom thinking. *Educational Psychology*, *13*(3-4), 291-304.
- Lipman, M. (1996). Philosophical discussion plans and exercises. *Analytic Teaching*, *16*(2).
- Lipman, M., & Sharp, A. M. (1978). Some educational presuppositions of philosophy for children. *Oxford review of Education*, *4*(1), 85-90.
- Lipman, M. (1988). Philosophy for children and critical thinking. *Thinking: The Journal of Philosophy for Children*, *7*(4), 40-42.
- Lipman, M. (1997). Thinking in community. *Inquiry: Critical thinking across the disciplines*, *16*(4), 6-21.
- Lipman, M. (1985). Philosophy and the Cultivation of Reasoning. *Thinking: The Journal of Philosophy for Children*, *5*(4), 33-41.
- Lipman, M. (1995). Good thinking. *Inquiry: Critical Thinking Across the Disciplines*, *15*(2), 37-41.
- Lipman, M. (1988). Critical thinking and the use of criteria. *Inquiry: Critical Thinking across the Disciplines*, *1*(2), 2-2.
- Lipman, M. (2011). Philosophy for children: Some assumptions and implications. *Ethics in Progress*, *2*(1), 3-16.
- Lipman, M. (1954). The relation of critical functions and critical decisions to art inquiry. *The Journal of Philosophy*, *51*(22), 653-667.
- Lipman, M. (1987). Preparing teachers to teach for thinking. *Philosophy Today*, *31*(1), 90-96.
- Lipman, M. (1992). Unreasonable people and inappropriate judgments. *Inquiry: Critical Thinking Across the Disciplines*, *10*(3), 1-1.

- Lipman, M. (1992). Criteria and judgment in critical thinking. *Inquiry: Critical Thinking Across the Disciplines*, *9*(4), 3-4.
- Lipman, M. (1982). Why Aren't Thinking Skills Being Taught? *Thinking: The Journal of Philosophy for Children*, *3*(3/4), 45-46.

Richard Nisbett

- Lehman, D. R., & Nisbett, R. E. (1990). A longitudinal study of the effects of undergraduate training on reasoning. *Developmental Psychology*, *26*(6), 952.
- Norenzayan, A., Smith, E. E., Kim, B. J., & Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. *Cognitive science*, *26*(5), 653-684.
- Nisbett, R. E. (2013). *Rules for reasoning*. Psychology Press.
- Fong, G. T., & Nisbett, R. E. (1991). Immediate and delayed transfer of training effects in statistical reasoning. *Journal of Experimental Psychology: General*, *120*(1), 34.
- Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive psychology*, 18(3), 253-292.
- E. Nisbett, R., & Norenzayan, A. (2002). *Culture and cognition. Stevens' handbook of experimental psychology*.
- Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science*, 238(4827), 625-631.

Richard Paul, Linda Elder

- Elder, L., & Paul, R. (2012). Critical Thinking: Competency Standards Essential to the Cultivation of Intellectual Skills, Part 4. *Journal of Developmental Education*, 35(3), 30–31.
- Elder, L., & Paul, R. (1994). Critical thinking: Why we must transform our teaching. *Journal of Developmental Education*, 18(1), 34.
- Elder, L., & Paul, R. (2008). Critical thinking: Strategies for improving student learning. *Journal of Developmental Education*, 32(1), 32.
- Elder, L., & Paul, R. (2009). *The aspiring thinker's guide to critical thinking*.
- Elder, L., & Paul, R. (2019). *The thinker's guide to analytic thinking: How to take thinking apart and what to look for when you do*. Rowman & Littlefield.
- Paul, R. (1981). Teaching critical thinking in the "strong" sense: A focus on self-deception, world views, and a dialectical mode of analysis. *Informal Logic*, 4(2).
- Paul, R. W. (1984). Critical thinking: fundamental to education for a free society. *Educational leadership*, 42(1), 4-14.
- Paul, R. W. (1985). Bloom's Taxonomy and Critical Thinking Instruction. *Educational leadership*, 42(8), 36-39.
- Paul, R. W. (1985). The critical-thinking movement. In *National Forum* (Vol. 65, No. 1, p. 2). Honor Society of Phi Kappa Phi.
- Paul, R. (1989). *Critical Thinking Handbook: High School. A Guide for Redesigning Instruction*. Center for Critical Thinking and Moral Critique, Sonoma State University, Rohnert Park, CA 94928.
- Paul, R. W., & Binker, A. J. A. (1990). *Critical thinking: What every person needs to survive in a rapidly changing world*. Center for Critical Thinking and Moral Critique, Sonoma State University, Rohnert Park, CA 94928.
- Paul, R. (1990). Critical thinking. Rohnert Park, CA: Sonoma State University.

- Paul, R., & Nosich, G. M. (1992). *A Model for the National Assessment of Higher Order Thinking*.
- Paul, R. W. (1993). The logic of creative and critical thinking. *American Behavioral Scientist*, 37(1), 21-39.
- Elder, L., & Paul, R. (1997). Critical thinking: Crucial distinctions for questioning. *Journal of Developmental Education*, *21*(2), 34.
- Paul, R. (2000). *Critical thinking, moral integrity and citizenship: Teaching for the intellectual virtues. Knowledge, belief and character: Readings in virtue epistemology*, 163-175.
- Paul, R., & Elder, L. (2003). Critical thinking: Teaching students how to study and learn (Part III). *Journal of Developmental Education*, 26(3), 36.
- Paul, R. (2005). *The state of critical thinking today. New directions for community colleges*, 2005(130), 27-38.
- Paul, R., & Elder, L. (2006). Critical thinking: The nature of critical and creative thought. *Journal of Developmental Education*, 30(2), 34.
- Paul, R., & Elder, L. (2007). Critical thinking: The art of Socratic questioning. *Journal of developmental education*, 31(1), 36.
- Paul, R., & Elder, L. (2013). *Critical thinking: Tools for taking charge of your professional and personal life*. Pearson Education.
- Paul, R. W. (2018). Critical thinking and the critical person. In *Thinking* (pp. 373-403). Routledge.
- Paul, R., & Elder, L. (2019). *A guide for educators to critical thinking competency standards: Standards, principles, performance indicators, and outcomes with a critical thinking master rubric*. Rowman & Littlefield.
- Paul, R., & Elder, L. (2019). *The nature and functions of critical & creative thinking*. Rowman & Littlefield.
- Walsh, D., & Paul, R. W. (1986). *The Goal of Critical Thinking: from Educational Ideal to Educational Reality*.

Keith Stanovich

- Stanovich, K. E., Sá, W. C., & West, R. F. (2004). Individual differences in thinking, reasoning, and decision-making. *The nature of reasoning*, 375-409.
- Stanovich, K. E., Toplak, M. E., & West, R. F. (2008). The development of rational thought: A taxonomy of heuristics and biases. In *Advances in child development and behavior* (Vol. 36, pp. 251-285). JAI.
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & cognition*, *39*(7), 1275.
- Sá, W. C., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of educational psychology*, *91*(3), 497.
- Kokis, J. V., Macpherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology*, *83*(1), 26-52.

Tim van Gelder

- Van Gelder, T. (2001, December). How to improve critical thinking using educational technology. In *Meeting at the crossroads: Proceedings of the 18th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education* (pp. 539-548).

Else

- Byrnes, J. P., & Dunbar, K. N. (2014). The nature and development of critical-analytic thinking. *Educational Psychology Review*, *26*(4), 477-493.
- Dwyer, C. P., Hogan, M. J., & Stewart, I. (2014). An integrated critical thinking framework for the 21st century. *Thinking Skills and Creativity*, *12*, 43-52.
- Hitchcock, D. (2017). Do the fallacies have a place in the teaching of reasoning skills or critical thinking? In *On reasoning and argument* (pp. 401-408). Springer, Cham
- Lilienfeld, S. O., Ammirati, R., & Landfield, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives on psychological science*, *4*(4), 390-398.
- McPeck, J. E. (2016). *Critical thinking and education*. Routledge.
- Siegel, H. (1991). The generalizability of critical thinking. *Educational Philosophy and Theory*, *23*(1), 18-30.
- Siegel, H. (1980, November). Critical thinking as an educational ideal. In *The Educational Forum* (Vol. 45, No. 1, pp. 7-23). Taylor & Francis Group.
- Siegel, H. (2013). *Educating reason*. Routledge.

General

- Baron, J. (1993). Why Teach Thinking? An Essay. *Applied Psychology*, 42(3), 191–214.
- Boisvert, J. (2015). Pensée critique: définition, illustration et applications. *Revue Québécoise de Psychologie*, Volume 36 Numéro 1, Pages 3-33.
- Pithers, R. T., & Soden, R. (2000). Critical thinking in education: A review. *Educational research*, *42*(3), 237-249.

3. Cognitive bases

- Acerbi, A. (2019a). Cognitive attraction and online misinformation. *Palgrave Communications*, *5*(1), 1-7.
- Acerbi, A. (2019b). *Cultural Evolution in the Digital Age*. Oxford University Press.
- Allom, V., Mullan, B., & Hagger, M. (2016). Does inhibitory control training improve health behaviour? A meta-analysis. *Health Psychology Review*, *10*(2), 168-186.
- Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states?. *Psychological review*, *116*(4), 953.
- Ariely, D. (2008). *Predictably irrational*. New York, NY: Harper Audio.
- Aron, A. R. (2007). The neural basis of inhibition in cognitive control. *The neuroscientist*, *13*(3), 214-228.
- Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: one decade on. *Trends in cognitive sciences*, *18*(4), 177-185.

- Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. *Organizational influence processes*, 295-303.
- Atran, S. (2002). *In gods we trust: The evolutionary landscape of religion*. Oxford University Press.
- Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90-109.
- Bago, B., & De Neys, W. (2019). The smart System 1: Evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking & Reasoning*, *25*(3), 257-299.
- Ball, LJ, Thompson, VA, & Stupple, EJN (2017). Conflict and dual process theory: The case of belief bias. In W. De Neys (Ed.), Dual Process Theory 2.0 (pp. 100-120). Abingdon, Oxon: Routledge. Conflict and Dual-Process Theory: The Case of Belief Bias.
- Battersby, M., & Bailin, S. (2013). Critical thinking and cognitive biases.
- Bebbington, K., MacLeod, C., Ellison, T. M., & Fay, N. (2017). The sky is falling: evidence of a negativity bias in the social transmission of information. *Evolution and Human Behavior*, *38*(1), 92-101.
- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*(5304), 1293-1295.
- Białek, M., & De Neys, W. (2017). Dual processes and moral conflict: Evidence for deontological reasoners' intuitive utilitarian sensitivity. *Judgment and Decision-making*, *12*(2), 148-167.
- Blaine, T., & Boyer, P. (2018). Origins of sinister rumors: A preference for threat-related material in the supply and demand of information. *Evolution and Human Behavior*, *39*(1), 67-75.
- Boyd, R., & Richerson, P. J. (2005). *The origin and evolution of cultures*. Oxford University Press.
- Boyer, P. (2018). *Minds make societies: How cognition explains the world humans create*. Yale University Press.
- Butler, A. C., Karpicke, J. D., & Roediger III, H. L. (2008). Correcting a metacognitive error: feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 918.
- Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning*, *1*(1), 69-84.
- Carpenter, J., Sherman, M. T., Kievit, R. A., Seth, A. K., Lau, H., & Fleming, S. M. (2019). Domain-general enhancements of metacognitive ability through adaptive training. *Journal of Experimental Psychology: General*, *148*(1), 51.
- Casscells, W., Schoenberger, A., & Graboys, T. B. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine*, *299*(18), 999-1001.
- Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: A quantitative approach* (No. 16). Princeton University Press.
- Chi, M., Glaser, R., & Rees, E. (1982). Experience and problem solving. *Advances in the psychology of human intelligence*, *1*.
- Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in cognitive sciences*, *15*(5), 218-226.

- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, *31*(3), 187-276.
- Cosmides, L., & Tooby, J. (1994). Better than rational : Evolutionary psychology and the invisible hand. *The American Economic Review*,*84*(2), 327-332.
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *cognition*, *58*(1), 1-73.
- De Neys, W., Vartanian, O., & Goel, V. (2008). Smarter than we think: When our brains detect that we are biased. *Psychological Science*, *19*(5), 483-489.
- De Neys, W., Moyens, E., & Vansteenwegen, D. (2010). Feeling we're biased: Autonomic arousal and reasoning conflict. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 208-216.
- De Neys, W., Rossi, S., & Houdé, O. (2013). Bats, balls, and substitution sensitivity: Cognitive misers are no happy fools. *Psychonomic Bulletin & Review*, *20*(2), 269-273.
- De Neys, W., Lubin, A., & Houdé, O. (2014). The smart nonconserver: Preschoolers detect their number conservation errors. *Child Development Research*, *2014*.
- Dawkins, R. (1982). Replicators and vehicles. *Current problems in sociobiology*, *45*, 64.
- Dawkins, R. (2016). *The selfish gene*. Oxford university press.
- Diamond, A. (1991). Neuropsychological insights into the meaning of object concept development.
- Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental cognitive neuroscience*, *18*, 34-48.
- Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. Sage Publications.
- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current directions in psychological science*, *12*(3), 83-87.
- Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, *13*(4), 277-287.
- Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational psychologist*, *46*(1), 6-25.
- Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational behavior and human decision processes*, *105*(1), 98-121.
- Enge, S., Behnke, A., Fleischhauer, M., Küttler, L., Kliegel, M., & Strobel, A. (2014). No evidence for true training and transfer effects after inhibitory control training in young healthy adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 987.
- Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science*, *8*(3), 223-241.
- Fischhoff, B. (1982). Debiasing. Judgment under uncertainty: Heuristics and biases. *Judgment under uncertainty: Heuristics and biases*, 422-444.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, *34*(10), 906.

- Fleming, S. M., & Daw, N. D. (2017). Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. Psychological Review, 124(1), 91.
- Fleming, S. M., Dolan, R. J., & Frith, C. D. (2012). Metacognition: computation, biology and function. The Royal Society.
- Fleming, S. M., & Frith, C. D. (Eds.). (2014). *The cognitive neuroscience of metacognition*. London, UK: Springer.
- Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. Frontiers in Human Neuroscience, 8, 443.
- Frederick, S. (2005). Cognitive reflection and decision-making. *Journal of Economic perspectives*, *19*(4), 25-42.
- Galef Jr, B. G., Dudley, K. E., & Whiskin, E. E. (2008). Social learning of food preferences in 'dissatisfied'and 'uncertain'Norway rats. *Animal Behaviour*, *75*(2), 631-637.
- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". *European review of social psychology*, *2*(1), 83-115.
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky.
- Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. *Blackwell handbook of judgment and decision-making*, *62*, 88.
- Gigerenzer, G., Hoffrage, U., & Goldstein, D. G. (2008). *Fast and frugal heuristics are plausible models of cognition: Reply to Dougherty*, Franco-Watkins, and Thomas (2008).
- Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition*, *43*(2), 127-171.
- Gigerenzer, G., & Todd, P. M. the ABC Research Group 1999. *Simple heuristics that make us smart*.
- Gilovich, T., & Griffin, D. (2002). Introduction-heuristics and biases: Then and now. *Heuristics and biases: The psychology of intuitive judgment*, 1-18.
- Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge university press.
- Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental psychology*, *37*(5), 620.
- Goupil, L., & Kouider, S. (2016). Behavioral and neural indices of metacognitive sensitivity in preverbal infants. Current Biology, 26(22), 3038–3045.
- Goupil, L., & Kouider, S. (2019). Developing a Reflective Mind: From Core Metacognition to Explicit Self-Reflection. Current Directions in Psychological Science, 0963721419848672.
- Goupil, L., Romand-Monnier, M., & Kouider, S. (2016). Infants ask for help when they know they don't know. *Proceedings of the National Academy of Sciences*, *113*(13), 3492-3496.
- Grimaldi, P., Lau, H., & Basso, M. A. (2015). There are things that we know that we know, and there are things that we do not know we do not know: Confidence in decision-making. Neuroscience & Biobehavioral Reviews, 55, 88–97.

- Halpern, D. F. (2013). *Thought and knowledge: An introduction to critical thinking*. Psychology Press.
- Harris, P. L. (2012). Trusting what you're told: How children learn from others. Harvard University Press.
- Harris, P. L., & Corriveau, K. H. (2011). Young children's selective trust in informants. Philosophical Transactions of the Royal Society B: Biological Sciences, 366(1567), 1179–1187.
- Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W. E., & Moore, T. (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition*, *27*(5), 733-763.
- Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and social psychology Review*, *10*(1), 47-66.
- Haselton, M. G., Nettle, D., & Murray, D. R. (2015). The evolution of cognitive bias. *The handbook of evolutionary psychology*, 1-20.
- Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of educational research*, *66*(2), 99-136.
- Heath, C., Bell, C., & Sternberg, E. (2001). Emotional selection in memes: the case of urban legends. *Journal of personality and social psychology*, *81*(6), 1028.
- Heilbron, M., & Meyniel, F. (2019). Confidence resets reveal hierarchical adaptive learning in humans. *PLoS computational biology*, *15*(4), e1006972.
- Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and human behavior*, *22*(3), 165-196.
- Higgins, S., Katsipataki, M., Villanueva-Aguilera, A. B., Coleman, R., Henderson, P., Major, L. E., ... & Mason, D. (2016). The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit.
- Houdé, O., & Borst, G. (2015). Evidence for an inhibitory-control theory of the reasoning brain. *Frontiers in human neuroscience*, *9*, 148.
- Hutchins, E. (1995). *Cognition in the Wild* (No. 1995). MIT press.
- Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive science*, *19*(3), 265-288.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Kahneman, D., Slovic, S. P., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge university press.
- Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions.
- Kepecs, A., & Mainen, Z. F. (2012). A computational framework for the study of confidence in humans and animals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1322-1337.
- Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, *455*(7210), 227-231.
- Klahr, D., Fay, A. L., & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive psychology*, *25*(1), 111-146.
- Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences*, *27*(12), 712-719.

- Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological review*, *100*(4), 609.
- Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological review*, *119*(1), 80.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, *77*(6), 1121.
- Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological review*, *96*(4), 674.
- Kuhn, D. (2005). *Education for thinking*. Harvard University Press.
- Kuhn, D. (2011). What is scientific thinking and how does it develop?.
- Kuhn, D., Garcia-Mila, M., Zohar, A., Andersen, C., White, S. H., Klahr, D., & Carver, S. M. (1995). Strategies of knowledge acquisition. *Monographs of the society for research in child development*, i-157.
- Jiménez, Á. V., & Mesoudi, A. (2019). Prestige-biased social learning: current evidence and outstanding questions. *Palgrave Communications*, *5*(1), 1-12.
- Lea, A. M., & Ryan, M. J. (2015). Irrationality in mate choice revealed by túngara frogs. *Science*, *349*(6251), 964-966.
- Lhermitte, F. (1983). 'Utilization behaviour'and its relation to lesions of the frontal lobes. *Brain*, *106*(2), 237-255.
- Lichtenstein, S. B. Fischoff, and L. Phillips (1982),"Calibration of Probabilities: The State of the Art to 1980,". *Judgement under Uncertainty: Heuristics and Biases*, 306-334.
- Lsockl, K., & Schneider, W. (2007). Knowledge about the mind: Links between theory of mind and later metamemory. *Child Development*, *78*(1), 148-167.
- Mascaro, O., & Morin, O. (2015). Epistemology for beginners: Two-to five-year-old children's representation of falsity. *PloS one*, *10*(10).
- Mercier, H. (2017). How gullible are we? A review of the evidence from psychology and social science. *Review of General Psychology*, *21*(2), 103-122.
- Mercier, H. (2020). *Not born yesterday: The science of who we trust and what we believe*. Princeton University Press.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory.
- Metcalfe, J., & Shimamura, A. P. (Eds.). (1994). *Metacognition: Knowing about knowing*. MIT press.
- Meyniel, F., Sigman, M., & Mainen, Z. F. (2015). Confidence as Bayesian probability: From neural origins to behavior. *Neuron*, *88*(1), 78-92
- Meyniel, F. A bayesian approach to the brain. dana.org/article/a-bayesian-approach-to-the-brain/
- Meyniel, F., & Dehaene, S. (2017). Brain networks for confidence weighting and hierarchical inference during probabilistic learning. *Proceedings of the National Academy of Sciences*, *114*(19), E3859-E3868.
- Middlebrooks, P. G., & Sommer, M. A. (2012). Neuronal correlates of metacognition in primate frontal cortex. *Neuron*, *75*(3), 517-530.
- Milgram, S. (1974). The dilemma of obedience. *The Phi Delta Kappan*, *55*(9), 603-606.

- Morgan, T. J., Laland, K. N., & Harris, P. L. (2015). The development of adaptive conformity in young children: effects of uncertainty and consensus. *Developmental science*, *18*(4), 511-524.
- Morris, B. J., Croker, S., Masnick, A. M., & Zimmerman, C. (2012). The emergence of scientific reasoning. In *Current topics in children's learning and cognition*. IntechOpen.
- Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing decisions: Improved decision-making with a single training intervention. *Policy Insights from the Behavioral and Brain Sciences*, *2*(1), 129-140.
- Neill, W. T., Valdes, L. A., & Terry, K. M. (1995). Selective attention and the inhibitory control of cognition. In *Interference and inhibition in cognition* (pp. 207-261). Academic Press.
- Newman, I. R., & Thompson, V. A. (2017). Logical intuitions and other conundra for dual process theories. In *Dual Process Theory 2.0* (pp. 129-144). Routledge.
- Nisbett, R. E. (2015). *Mindware: Tools for smart thinking*. Farrar, Straus and Giroux.
- Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science*, *238*(4827), 625-631.
- Norman, E., Pfuhl, G., Sæle, R. G., Svartdal, F., Låg, T., & Dahl, T. I. (2019). Metacognition in psychology. *Review of General Psychology*, *23*(4), 403-424.
- Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*(4), 608.
- Ott, T., Masset, P., & Kepecs, A. (2018, January). The neurobiology of confidence: from beliefs to neurons. In *Cold Spring Harbor symposia on quantitative biology* (Vol. 83, pp. 9-16). Cold Spring Harbor Laboratory Press.
- Palmer, E. C., David, A. S., & Fleming, S. M. (2014). Effects of age on metacognitive efficiency. *Consciousness and cognition*, *28*, 151-160.
- Pasquinelli, E. (2006). *An analysis of the notion of illusion and illusory phenomena. Illusions in haptic, dynamic, kinesthetic touch* (Doctoral dissertation).
- Pasquinelli, E. (2012). The awareness of illusions. In *Perceptual Illusions* (pp. 59-74). Palgrave Macmillan, London.
- Pennycook, G. (2017). A perspective on the theoretical foundation of dual-process models. *Dual process theory*, *2*, 34.
- Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). Everyday consequences of analytic thinking. *Current Directions in Psychological Science*, *24*(6), 425-432.
- Pennycook, G., De Neys, W., Evans, J. S. B., Stanovich, K. E., & Thompson, V. A. (2018). The mythical dual-process typology. *Trends in Cognitive Sciences*, *22*(8), 667-668.
- Pennycook, G., Ross, R. M., Koehler, D. J., & Fugelsang, J. A. (2017). Dunning–Kruger effects in reasoning: Theoretical implications of the failure to recognize incompetence. *Psychonomic Bulletin & Review*, *24*(6), 1774-1784.
- Pohl, R. (2017). Cognitive illusions: Intriguing phenomena in judgment, thinking, and memory. *Abingdon, UK: Routledge*.
- Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: knowns and unknowns. *Nature neuroscience*, *16*(9), 1170.
- Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty: distinct probabilistic quantities for different goals. *Nature neuroscience*, *19*(3), 366.

- Proust, J. (2013). *The philosophy of metacognition: Mental agency and self-awareness*. OUP Oxford.
- Proust, J. La métacognition. Les enjeux pédagogiques de la recherche. https://educamus.ac-versailles.fr/IMG/pdf/metacognition_gt5.pdf
- Rakoczy, H. (2017). In defense of a developmental dogma: Children acquire propositional attitude folk psychology around age 4. *Synthese*, *194*(3), 689-707.
- Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, *2*(1), 79-87.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, *2*, 64-99.
- Richerson, P. J., & Boyd, R. (2008). *Not by genes alone: How culture transformed human evolution*. University of Chicago press.
- Rode, C., Cosmides, L., Hell, W., & Tooby, J. (1999). When and why do people avoid unknown probabilities in decisions under uncertainty? Testing some predictions from optimal foraging theory. *Cognition*, *72*(3), 269-304.
- Rouault, M., McWilliams, A., Allen, M. G., & Fleming, S. M. (2018). Human metacognition across domains: insights from individual differences and neuroimaging. *Personality neuroscience*, *1*.
- Rouault, M., Dayan, P., & Fleming, S. M. (2019). Forming global estimates of self-performance from local confidence. *Nature communications*, *10*(1), 1-11.
- Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive science*, *26*(5), 521-562.
- Shadlen, M. N., & Kiani, R. (2013). Decision-making as a window on cognition. *Neuron*, *80*(3), 791-806.
- Sharot, T. (2011). The optimism bias. *Current biology*, *21*(23), R941-R945.
- Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature neuroscience*, *14*(11), 1475.
- Sharot, T., Kanai, R., Marston, D., Korn, C. W., Rees, G., & Dolan, R. J. (2012). Selectively altering belief formation in the human brain. *Proceedings of the National Academy of Sciences*, *109*(42), 17058-17062.
- Sharot, T., & Garrett, N. (2016). Forming beliefs: Why valence matters. *Trends in cognitive sciences*, *20*(1), 25-33.
- Schauble, L. (1990). Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of experimental child psychology*, *49*(1), 31-57.
- Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, *32*(1), 102.
- Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., & Frith, C. D. (2014). Supra-personal cognitive control and metacognition. *Trends in cognitive sciences*, *18*(4), 186-193.
- Schulz, L. E., & Gopnik, A. (2004). Causal learning across domains. *Developmental psychology*, *40*(2), 162.
- Schulz, L. E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental science*, *10*(3), 322-332.

- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593-1599.
- Serra, M. J., & Metcalfe, J. (2009). 15 Effective Implementation of Metacognition. *Handbook of metacognition in education*, 278.
- Scott, R. M., & Baillargeon, R. (2017). Early false-belief understanding. *Trends in Cognitive Sciences*, *21*(4), 237-249.
- Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. (2016). Do "brain-training" programs work?. *Psychological Science in the Public Interest*, *17*(3), 103-186.
- Sloman, S., & Fernbach, P. (2018). *The knowledge illusion: Why we never think alone*. Penguin.
- Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and brain sciences*, *26*(3), 317-339.
- Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental science*, *10*(1), 89-96.
- Sperber, D. (1996). Explaining culture: A naturalistic approach. *Cambridge, MA: Cambridge*.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. Mind & Language, 25(4), 359–393.
- Stanovich, K. E., & Stanovich, P. J. (2010). *A framework for critical thinking, rational thinking, and intelligence. Innovations in educational psychology: Perspectives on learning, teaching and human development*, 195-237.
- Stephens, D. W. (2008). Decision ecology: foraging and the ecology of animal decision-making. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(4), 475-484.
- Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning* (Vol. 135). Cambridge: MIT press.
- Tamis-LeMonda, C. S., Adolph, K. E., Lobo, S. A., Karasik, L. B., Ishak, S., & Dimitropoulou, K. A. (2008). When infants take mothers' advice: 18-month-olds integrate perceptual and social information to guide motor action. *Developmental psychology*, *44*(3), 734.
- Tipper, S. P. (2001). Does negative priming reflect inhibitory mechanisms? A review and integration of conflicting views. *The Quarterly Journal of Experimental Psychology Section A*, *54*(2), 321-343.
- Trippas, D., & Handley, S. J. (2018). The parallel processing model of belief bias: Review and extensions.
- Trouche, E., Sander, E., & Mercier, H. (2014). Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General*, *143*(5), 1958.
- Trouche, E., Shao, J., & Mercier, H. (2019). Objective evaluation of demonstrative arguments. *Argumentation*, *33*(1), 23-43.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124-1131.
- van Bergen, Y., Coolen, I., & Laland, K. N. (2004). Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1542), 957-962.

- Vo, V. A., Li, R., Kornell, N., Pouget, A., & Cantlon, J. F. (2014). Young children bet on their numerical skills: metacognition in the numerical domain. *Psychological science*, *25*(9), 1712-1721.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child development*, *72*(3), 655-684
- Yaniv, I. (2004). Receiving other people's advice: Influence and benefit. *Organizational behavior and human decision processes*, *93*(1), 1-13.
- Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1310-1321.
- Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental review*, *20*(1), 99-149.
- Zimmerman, C., Raghavan, K., & Sartoris, M. (2003). The impact of the MARS curriculum on students' ability to coordinate theory and evidence. *International Journal of Science Education*, *25*(10), 1247-1271.

4. Educability of CT

- Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. *Review of Educational Research*, *85*(2), 275-314.
- Adey, P., Shayer, M., & Yates, C. (1989). Thinking science: Student and teachers' materials for the CASE intervention. *London, Mccacmillan*.
- Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. *Educational researcher*, *25*(4), 5-11.
- Bassok, M., & Olseth, K. L. (1995). Judging a book by its cover: Interpretative effects of content on problem-solving transfer. *Memory & Cognition*, *23*(3), 354-367.
- Biederman, I., & Shiffrar, M. M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. *Journal of Experimental Psychology: Learning, memory, and cognition*, *13*(4), 640.
- Bjorklund, D. F., & Harnishfeger, K. K. (1990). Children's strategies: Their definition and origins. *Children's strategies: Contemporary views of cognitive development*, 309-323.
- Blagg, N. (1988). *SOMERSET Thinking Skills Course. Handbook + Modules 1-5: 1. Foundations for Problem Solving. 2. Analysing and Synthesising. 3. Comparative Thinking. 4. Positions in Time and Space. 5. Understanding Analogies*. Blackwell/Somerset County Council.
- Boku, M., Yama, H., & Mercier, H. (2018). Robust improvements on reasoning performance following discussion in Japan. *Japanese Psychological Research*, *60*(1), 47-53.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn* (Vol. 11). Washington, DC: National academy press.
- Brown, A. L., Palincsar, A. S., & Armbruster, B. B. (1984). Instructing comprehension-fostering activities in interactive learning situations. *Learning and comprehension of text*, 255-286.

- Brown, P. C., Roediger III, H. L., & McDaniel, M. A. (2014). *Make it stick*. Harvard University Press.
- Carraher, T. N. (1986). From drawings to buildings; working with mathematical scales. *International Journal of Behavioral Development*, *9*(4), 527-544.
- Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. *The psychology of learning and motivation*, *16*, 1-58.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55–81.
- Chi, M. T. H., & Feltovich, P. R. Glaser, and R. Rees. (1982). Expertise in problem solving. *Advances in the psychology of human intelligence*, *1*, 1-75.
- Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive science*, *5*(2), 121-152.
- Chi, M. T., Glaser, R., & Farr, M. J. (1988). The nature of expertise. Hillsdale, NJ: L.
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. *The adapted mind: Evolutionary psychology and the generation of culture*, *163*, 163-228.
- de Groot, A. D. (1965). *Thought and choice in chess* (2nd ed. 1978). The Hague: Mouton.
- Dunker, K. (1945). On problem solving. Psych. *Monographs*, *58*, 270.
- Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational researcher*, *18*(3), 4-10.
- Ericsson, K. A., & Chase, W. G. (1982). *Exceptional memory* (No. TR-8-ONR). Carnegie-Mellon Univ Pittsburgh Pa Dept Of Psychology.
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, *100*(3), 363.
- Facione, P. A. (1991). Using the California Critical Thinking Skills Test in Research, Evaluation, and Assessment.
- Facione, P. A. (2000). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. *Informal logic*, *20*(1).
- Fayol, M., & Monteil, J. M. (1994). Note de synthèse. *Revue française de pédagogie*, *106*(1), 91-110.
- Feuerstein, R., Ya'acov Rand, Hoffman, M. B., & Miller, R. (1980). Instrumental enrichment: An intervention program for cognitive modifiability.
- Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive psychology*, *18*(3), 253-29
- Gamo, S., Sander, E., & Richard, J. F. (2010). Transfer of strategy use by semantic recoding in arithmetic problem solving. *Learning and Instruction*, *20*(5), 400-410.
- Gelder, T. V. (2005). Teaching critical thinking: Some lessons from cognitive science. *College teaching*, *53*(1), 41-48.
- Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, *95*(2), 393.
- Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive psychology*, *12*(3), 306-355.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive psychology*, *15*(1), 1-38.
- Gobet, F., & Simon, H. A. (1996). Recall of random and distorted chess positions: Implications for the theory of expertise. *Memory & cognition*, *24*(4), 493-503.

- Halpern, D. F. (2013). *Thought and knowledge: An introduction to critical thinking*. Psychology Press.
- Hattie, J. A., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. *npj Science of Learning*, *1*, 16013.
- Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of educational research*, *66*(2), 99-136.
- Helfenstein, S., & Saariluoma, P. (2006). Mental contents in transfer. *Psychological Research*, *70*(4), 293-303.
- Helsdingen, A. S., Van den Bosch, K., Van Gog, T., & van Merriënboer, J. J. (2010). The effects of critical thinking instruction on training complex decision-making. *Human factors*, *52*(4), 537-545.
- Hernstein, R. J., Nickerson, R. S., de Sanchez, M., & Swets, J. A. (1986). Teaching thinking skills. *American Psychologist*, *41*(11), 1279.
- Higgins, S., Hall, E., Baumfield, V., & Moseley, D. (2005). A meta-analysis of the impact of the implementation of thinking skills approaches on pupils.
- Higgins, S., Katsipataki, M., Villanueva-Aguilera, A. B., Coleman, R., Henderson, P., Major, L. E., ... & Mason, D. (2016). The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit.
- Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & cognition*, *15*(4), 332-340.
- Kosonen, P., & Winne, P. H. (1995). Effects of teaching statistical laws on reasoning about everyday problems. *Journal of Educational Psychology*, *87*(1), 33.
- Kuhn, D. (2015). Thinking together and alone. *Educational researcher*, *44*(1), 46-53.
- Kuhn, D., Hemberger, L., & Khait, V. (2016). Tracing the development of argumentive writing in a discourse-rich context. *Written Communication*, *33*(1), 92-121.
- Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational research review*, *10*, 133-149.
- Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*(4450), 1335-1342.
- Lehman, D. R., Lempert, R. O., & Nisbett, R. E. (1988). The effects of graduate training on reasoning: Formal discipline and thinking about everyday-life events. *American Psychologist*, *43*(6), 431.
- Lipman, M. S., & Sharp, A. M. AM and Oscanyan, FS (1980). Philosophy in the classroom.
- Luchins, A. S., & Luchins, E. H. (1970). The effects of order of presentation of information and explanatory models. *The Journal of Social Psychology*, *80*(1), 63-70.
- Marin, L. M., & Halpern, D. F. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains. *Thinking Skills and Creativity*, *6*(1), 1-13.
- Marzano, R. J. (1998). A theory-based meta-analysis of research on instruction.
- Mercier, H. (2016). The argumentative theory: Predictions and empirical evidence. *Trends in Cognitive Sciences*, *20*(9), 689-700.
- Mercier, H., & Sperber, D. (2011). Argumentation: its adaptiveness and efficacy. *Behavioral and Brain Sciences*, *34*(2), 94.

- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.
- Minervino, R. A., Olguín, V., & Trench, M. (2017). Promoting interdomain analogical transfer: When creating a problem helps to solve a problem. *Memory & cognition*, *45*(2), 221-232.
- Nisbett, C. J., Jepson, R. E., & Krantz, D. H. (1993). Inductive reasoning: Competence or skill? *Rules for reasoning*, 70.
- Nisbett, R. E. (2013). *Rules for reasoning*. Psychology Press.
- Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science*, *238*(4827), 625-631.
- Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological review*, *90*(4), 339.
- Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, memory, and cognition*, *14*(3), 510.
- Perkins, D. N., & Salomon, G. (1992). Transfer of learning. *International encyclopedia of education*, *2*, 6452-6457.
- Ramus, F. (2012). Au-delà de l’inné et de l’acquis. *Hors-série La recherche-jeux*, 18-20.
- Raynal, L., Clément, E., & Sander, E. (2017). Challenging the superficial similarities superiority account for analogical retrieval.
- Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, *80*(1), 16.
- Resnick, L. B. (1987). *Education and learning to think*. National Academies.
- Sala, G., & Gobet, F. (2017). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current Directions in Psychological Science, 26*(6), 515-520.
- Salomon, G., & Perkins, D. (1988). Teaching for transfer. *Educational Leadership*, *46*(1), 22-32.
- Salomon, G., & Perkins, D. N. (1987). Transfer of cognitive skills from programming: When and how? *Journal of educational computing research*, *3*(2), 149-169.
- Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational psychologist*, *24*(2), 113-142.
- Sills, J., Rowse, G., & Emerson, L. M. (2016). The role of collaboration in the cognitive development of young children: a systematic review. *Child: care, health and development*, *42*(3), 313-324.
- Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill* (No. 9). Harvard University Press.
- Slavin, R. E. (2014). Cooperative Learning and Academic Achievement: Why Does Groupwork Work? [Aprendizaje cooperativo y rendimiento académico: ¿por qué funciona el trabajo en grupo?] *Anales de psicología/annals of psychology*, *30*(3), 785-791.
- Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions: III. Functions involving attention, observation and discrimination. *Psychological Review*, *8*(6), 553.
- Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. *The adapted mind: Evolutionary psychology and the generation of culture*, *19*.

- Trouche, E., Sander, E., & Mercier, H. (2014). Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General*, *143*(5), 1958.
- Van Gelder, T. (2001, December). How to improve critical thinking using educational technology. In *Meeting at the crossroads: Proceedings of the 18th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education* (pp. 539-548).
- Willingham, D. T. (2007). Critical thinking: Why it is so hard to teach? *American Federation of Teachers summer 2007, p. 8-19*.
- Willingham, D. (2019). How to teach critical thinking.
- Woodworth, R. S., & Thorndike, E. L. (1901). The influence of improvement in one mental function upon the efficiency of other functions.(I). *Psychological review*, *8*(3), 247.

Articles used by (Abrami et al. 2105) for their meta-analysis :

*****Abdulghani, B. A. (2003). *An inquiry into the effects of cooperative learning on critical thinking and achievement in the Arabic language by female high school students in the United Arab Emirates* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3114806)

*****Affeldt, J. E. O. (1994). *Occupational therapy clinical reasoning: supervision for the transition from student to clinician (critical thinking)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9411515)

*****Al-Khayyatt, A. K. A. (1980). *An experimental study comparing the effects of the inquiry method and the traditional method for teaching social studies in two Kuwaiti secondary schools for boys* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8107532)

*****Allison, R. D. (1972). *An investigation into the attitudes toward science of college chemistry students as a function of laboratory experience.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7300253)

*****Anderson, J. L. (1992). *A study of critical thinking as it relates to Myers-Briggs personality type, dominant function, and integrated process skills of ninth-grade physical science students.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9400005)

*****Angeli, C., & Valanides, N. (2009). Instructional effects on critical thinking: Performance on ill-defined issues. *Learning and Instruction, 19,* 322-334. http://dx.doi.org/10.1016/j.learninstruc.2008.06.010

*****Annis, D. B., & Annis, L. (1979). An empirical study of the impact of philosophy on students' CT ability. *Teaching Philosophy*, *3*, 145-152.

*****Apichatibutarapong, S., Logan, P., & Cobbin, D. (2004). Alternative strategies for teaching critical thinking. In J. Renner & J. Cross (Eds.), *New challenges in sustainability and growth in higher education: EDU-COM 2004 conference proceedings, 24-26 November 2004* (pp. 29-35). Joondalup, Western Australia: Edith Cowan University. Retrieved from

*****Arlo, R. (1968). *The relative effectiveness of inductive and expository teaching of principles of general semantics upon the critical reading ability of 9th-grade students.*

(Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6921178)

*****Armstrong, N. A. (1970). *The effect of two instructional inquiry strategies on critical thinking and achievement in eighth grade social studies* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7017235)

*****Arrufat, O. (1997). *The role of exploration and critical decision-making and problem solving in making life choices* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9813418)

*****Asamani, J. B. B. (1998). *An experimental study of the relationship between cooperative learning and critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9806590)

*****Aten, L. L. (2004). *Using guided learning journals to foster reflection and professional judgment in teachers* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3111581)

*****Athman, J. A. (2003). *The effects of environment-based education on students' critical thinking and achievement motivation* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3105582)

*****Ayoub, J. L. (1997). *The effect of classroom use of an electronic group support system on student critical thinking, performance and satisfaction* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9806760)

*****Bagby, L. A. (1991). *The study of popular culture to promote critical thinking in writing* (Master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1341542)

*****Bamvakais, K. M. (1981). *Promoting critical thinking: An instructional program for enhancing the remote transfer of critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8212317)

*****Bates, D. S. (1991). *Effect of a systematic inquiry intervention on the critical thinking skills of a sample of master's degree students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9206908)

*****Baumgartner, A. (1965). *An experimental study of speech fundamentals unit in public speaking upon organization skill and critical thinking ability* (Unpublished master's thesis). University of Nebraska, Lincoln.

*****Bean, A. T. (1969). *Some comparisons between conventional college teaching methods and a composite of procedures involving large lecture groups, seminars, and reduced class time* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6919928)

*****Belinsky, S. B. (2000). *The effect of a peer to peer strategy within radiation therapy and nursing clinical settings on the development of critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9961539)

*****Biddle, W. W. (1932). *Propaganda and education*. New York: AMS Press Inc.

*****Bird, J. J. (1984). *Effects of fifth graders' attitudes and critical thinking/reading skills resulting from a Junior Great Books program* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8411599)

*****Black, D. H. (1982). *Comparison of inquiry and non-inquiry methods in the development of critical thinking* (Doctoral dissertation). Available from ProQuest

Dissertations and Theses database. (UMI No. 8208436)
*****Bonnette, R., McBride, R. E., & Tolson, H. (2001). The differential effect of indirect instruction in the teaching of sport skills on critical thinking and self-esteem of early adolescent boys placed at risk. *Sport, Education & Society, 6*, 183-198. http://dx.doi.org/10.1080/13573320120084263
*****Bosma, B. A. (1981). *An experimental study to determine the feasibility of using folk literature to teach select critical reading skills to sixth graders* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8126472)
*****Boyadjian-Samawi, Z. (2006). *The effect of concept mapping on critical thinking skills and dispositions of junior and senior baccalaureate nursing students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3207567)
*****Brownell, J. A. (1953). The influence of training in reading in the social studies on the ability to think critically. *California Journal of Educational Research, 4*, 28-31.
*****Bruning, R., Siwatu, K. O., Liu, X. Y., PytlikZillig, L. M., Horn, C., Sic, S., & Carlson, D. (2008). Introducing teaching cases with face-to-face and computer-mediated discussion: Two multi-classroom quasi-experiments. *Contemporary Educational Psychology, 33*(2), 299-326. http://dx.doi.org/10.1016/j.cedpsych.2006.11.002
*****Burkhart, L. M. (2006). *Thinking critically about critical thinking: Developing thinking skills among high school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3214755)
*****Burns, E. (2009). *The use of science inquiry and its effect on critical thinking skills and dispositions in third grade students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3367974)
*****Butchart, S., Forster, D., Gold, I., Bigelow, J., Korb, K., Oppy, G., & Serrenti, A. (2009). Improving critical thinking using web based argument mapping exercises with automated feedback. *Australasian Journal of Educational Technology, 25*, 268-291. Retrieved from http://www.ascilite.org.au/ajet/ajet25/butchart.pdf
*****Butchart, S., Handfield, T., & Restall, G. (2009). Using peer instruction to teach philosophy, logic, and critical thinking. *Teaching Philosophy, 32*(1), 1-40.
*****Butler, G. P. (1990). *A study of the effects of utilizing the writing process as an instructional strategy on student thinking skills in ninth-grade biology courses* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9030670)
*****Byrne, J. S. (1983). *The effect of critical thinking skills instruction on achievement and attitudes of elementary students differing in learning style preferences* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8417843)
*****Cabrera, G. A. (1982). *Evaluating the teaching of critical thinking in content areas* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8301629)
*****Carter, L. M. (2008). Critical thinking dispositions in online nursing education. *Journal of Distance Education, 22*(3), 89-114. Retrieved from http://www.jofde.ca/index.php/jde/article/view/454/772
*****Carwie, L. (2009). *The effect of computer-supported argument mapping on the critical thinking skills of undergraduate nursing students* (Doctoral dissertation).

Available from ProQuest Dissertations and Theses database. (UMI No. 3388766)

*Chamberlain, M. A. (1993). *Philosophy for Children program and the development of critical thinking of gifted elementary students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9319950)

*Champion, B. W. (1975). *An inductive approach to the discrimination of factual and inferential statements as an aspect of critical reading* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7528097)

*Chen, F. C., & Lin, M. C. (2003). Effects of a nursing literature reading course on promoting critical thinking in two-year nursing program students. *Journal of Nursing Research, 11*, 137-147. http://dx.doi.org/10.1097/01.JNR.0000347629.80100.b2

*Chennault, A. (1989). *Enhancing critical thinking skills in gifted elementary school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8917052)

*Clark, R. C. (1977). *The effects of graphic materials on critical thinking in social studies* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7726650)

*Colbert, K. R. (1987). The effects of ceda and ndt debate training on critical thinking ability. *Journal of the American Forensic Association, 23*(4), 194-201.

*Colledge, D. G. (1993). *An experimental study comparing embedding and immersion approaches to instructing critical thinking to 5th-grade students (fifth-grade)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9334716)

*Comerford, S. S., Busk, P. L., & Roberts, W. L. (2000, April). *Epistemological beliefs of community-college students*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED442483)

*Conklin, K. E. (1986). *The impact of a critical thinking intervention on tenth grade students' expository writing skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8629807)

*Constantine, M. (1968). *An experimental study of the development of critical thinking skills of high school English teachers enrolled in a methods course* (Interim Report, Project No. HE-145). Illinois State-wide Curriculum Study Center in the Preparation of Secondary School English Teachers (ISCPET).

*Cook, J. E. (1972). *A study in critical listening using eight to ten year olds in an analysis of commercial propaganda emanating from television* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7226840)

*Crawford, J. A. (1976). *An investigation of the transfer of critical thinking skills from language arts to social studies* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7703254)

*Criner, L. A. (1992). *Teaching thinking and reasoning: A study of critical thinking in adults* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9237328)

*Crow, L. W., & Haws, S. G. (1985, April). *The effects of teaching logical reasoning upon students' critical thinking and science achievement*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, French Lick Springs, IN. (ERIC Document Reproduction Service No. ED255371)

*Curtis, C. K. (1980). Developing critical thinking skills in nonacademic social studies classes. *Alberta Journal of Educational Research*, *26*, 75-84.

*Dale, P. M., Ballotti, D., Handa, S., & Zych, T. (1997). An approach to teaching problem solving in the classroom. *College Student Journal*, *31*, 76-79.

*Daodee, S. (2002). *The implementation of teaching strategies for critical thinking development in nursing education in Thailand: A concerns-based approach* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. C809151)

*Desnouee, J. R. (2000). *The influence of the artforming technique on the self-esteem and reading behaviors of at-risk first-year college students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9980926)

*Dewey, J., & Bento, J. (2009). Activating children's thinking skills (ACTS): The effects of an infusion approach to teaching thinking in primary schools. *British Journal of Educational Psychology*, *79*, 329-351. http://dx.doi.org/10.1348/000709908X344754

*Diket, M. R. M. (1991). *Art criticism: Relationships to critical thinking, appreciation, and creativity among the gifted* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9206946)

*Dixon, F. A. (1996). *The use of the Hegelian dialectic as a heuristic in literature classes to improve critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9725536)

*Dolan, S. L. (1976). *The effect of value-conflict discussion on the critical thinking and reading of adults* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7626218)

*Donawa, A. (2009). *Critical thinking instruction and minority engineering students at a public urban higher education institution* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3396399)

*Douglas, J. (1951). An experimental study of training in problem solving methods. *Speech Monographs*, *19*, 115-116.

*Downing, K., Kwong, T., Chan, S.-W., Lam, T.-F., & Downing, W.-K. (2009). Problem-based learning and the development of metacognition. *Higher Education*, *57*, 609-621. http://dx.doi.org/10.1007/s10734-008-9165-x

*Drews, E. M. (1965). *Being and becoming--a cosmic approach to counseling and curriculum, part 2--the creative intellectual style in gifted adolescents*. East Lansing: Michigan State University.

*Duatepe-Paksu, A., & Ubuz, B. (2009). Effects of drama-based geometry instruction on student achievement, attitudes, and thinking levels. *Journal of Educational Research*, *102*, 272-286. http://dx.doi.org/10.3200/JOER.102.4.272-286

*Duckworth, J. B. (1968). *The effect of instruction in general semantics on the critical thinking of tenth and eleventh grade students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6906092)

*Dukess, S. L. (1986). *Teaching literary skills to eighth graders through critical thinking: A theoretical model and a comparison of two approaches* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8522954)

*Duphorne, P. L., & Gunawardena, C. N. (2005). The effect of three computer conferencing designs on critical thinking skills of nursing students. *American Journal of Distance Education*, *19*, 37-50. http://dx.doi.org/10.1207/s15389286ajde1901_3

*****Eason, L. E. (1986). *The relationship of critical thinking skills and psychological type in community college students' responses to science instruction* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8704157)

*****Eimers, S. P. (1987). *Effects of reflective thinking instruction taught through the philosophy for children program on the reasoning ability and attitude toward teaching thinking of preservice elementary teachers* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8809971)

*****Elliott, B. M. (1999). *The influence of an interdisciplinary course on critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9945800)

*****Endler, L. C., & Bond, T. G. (2008). Changing science outcomes: Cognitive acceleration in a US setting. *Research in Science Education*, *38*, 149-166. http://dx.doi.org/10.1007/s11165-007-9042-0

*****Ennis, T. R. (2007). *Assessing and enhancing critical thinking skills: Enhanced Peer Review (EPR) with the CAT instrument* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3263266)

*****Ersözlü, Z. N., & Arslan, M. (2009). The effect of developing reflective thinking on metacognitional awareness at primary education level in Turkey. *Reflective Practice, 10*, 683-695. http://dx.doi.org/10.1080/14623940903290752

*****Estrada, J. M. (1992). *Effects of programming in basic and logo on problem-solving skills* (Master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1349192)

*****Fang, R. J., Lin, C. C., Yang, H. J., Lee, C. J., Tsai, H. L., & Tsai, T. S. (2008). A study to increase the critical thinking and problem solving abilities by web-based learning. In Q. Li, S. Y. Chen, A. Xu & M. Li (Eds.), *Musp '08: Multimedia Systems and Signal Processing* (pp. 75-80).

*****Faryniarz, J. V. (1990). *The effectiveness of microcomputer simulators to stimulate environmental problem-solving with community college students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9002473)

*****Faryniarz, J. V., & Lockwood, L. G. (1992). Effectiveness of microcomputer simulations in stimulating environmental problem solving by community college students. *Journal of Research in Science Teaching, 29*, 453-470. http://dx.doi.org/10.1002/tea.3660290503

*****Fennelly, P. (1989). *The effects of one program in critical thinking instruction on the reading comprehension ability of underprepared college students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8813592)

*****Ferrell, S. D. (1992). *Critical thinking as a function of teacher questions* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9206949)

*****Feuerstein, M. (2002). *Media literacy in support of critical thinking* (Unpublished doctoral dissertation). Liverpool, England.

*****Fiasca, M. A. (1966). *Feasibility of integration of selected aspects of (CBA) Chemistry, (CHEMS) Chemistry and (PSSC) Physics into a two year Physical Science sequence*. Ann Arbor, MI. University Microfilms.

*****Fielding, R. (1985). A comparative study of instructional modes for developing visual art criticism skills. *Journal of the Institute of Art Education, 9*(3), 19-28.
*****Fischer, C. W. (1990). *Effects of a developmental drama-inquiry process on creative and critical thinking skills in early adolescent students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9010183)
*****Fisher, D. E. L. (1968). *An experimental study of the teaching of certain critical thinking skills to ninth grade pupils* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6814977)
*****Forbes, C. A. (1997). *Analyzing the growth of the critical thinking skills of college calculus students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9737709)
*****Frank, A. (1964). *An experiment in teaching critical thinking in a high school speech fundamentals course* (Unpublished doctoral dissertation). University of Wisconsin, Madison.
*****Frank, A. D. (1969). Teaching high school speech to improve critical thinking ability. *The Speech Teacher, 18*, 297-302.
*****Freeman, B. (1995). *A study of the efficacy of a developmental instruction design for non-traditional adult learners (nontraditional students)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9432514)
*****Friend, H. A. (1967). *A comparison of the relative effectiveness of two methods of teaching the course Time, Space, and Matter to selected eighth grade pupils* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6911745)
*****Gadzella, B. M., Hartsoe, K., & Harper, J. (1989). Critical thinking and mental ability groups. *Psychological Reports, 65*(3, Pt 1), 1019-1026.
*****Galinski, M. E. (1989). *An investigation of the relationship of microcomputer problem-solving software and critical thinking abilities of junior-high students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8818884)
*****Gent, P. J. (1997). *Using case studies in a course on disabilities to improve the attitudes toward inclusion and the critical thinking skills of preservice educators* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9716983)
*****Germann, P. J. (1989). Directed-inquiry approach to learning science process skills: treatment effects and aptitude-treatment interactions. *Journal of Research in Science Teaching, 26*, 237-250. http://dx.doi.org/10.1002/tea.3660260305
*****Ghee, H. J. (1976). *A study of the effects of high level cognitive questions on the levels of response and critical thinking abilities in students of two social problems classes* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7601090)
*****Gibbs, L. E., Browne, M. N., & Keeley, S. M. (1988). *Stimulating critical thinking through faculty development: design, evaluation, and problems*. Washington, DC: American Association of State Colleges and Universities, Wisconsin University.
*****Gibson, H. W. (1986). *Critical thinking: A communication model* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8601471)
*****Giles, K. S. (1990). *Improving students' transactional writing through cognitive*

strategy intervention (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9008808)

*****Gillespie, M. A. (2005). *Critical thinking about values: The effects of an instructional program, reasons for attending college, and general life goals on the application of critical thinking to values expressed in an essay prompt* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3193386)

*****Glaser, E. M. (1941). *An experiment in the development of critical thinking*. New York: AMS Press.

*****Griffin, C. M. (1979). *The effect of the use of selected newspaper articles in the teaching of certain critical reading/thinking skills* (Unpublished M.Ed. Thesis). Rutgers University.

*****Gudaitis, D. J. (1971). *The effects of two seventh grade science programs, "interaction of man and the biosphere" and "science is explaining", on student attitude, science processes, and critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7124823)

*****Guice, E. D. (1992). *The effect of instruction in concept analysis on critical thinking skills and moral reasoning decisions of senior baccalaureate nursing students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9225819)

*****Hagan, M. W. (2005). *Critical thinking skills study: Vocational nursing* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3174552)

*****Hagelskamp, J. L. (2001). *The effect of teaching critical thinking by infusion, with focus on transfer of skills, on the moral judgment and critical thinking of secondary school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9988873)

*****Hamilton, L. J. (1994). *Relationship of critical thinking and learning styles to nursing diagnosis* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9411644)

*****Hancock, B. W., Jr. (1982). *The effect of guided design on the critical thinking ability of college level administrative science students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8206458)

*****Hancock, B. W., Coscarelli, W. C., & White, G. P. (1983). Critical thinking and content acquisition using a modified guided design process for large course sections. *Educational and Psychological Research*, *3*(3), 139-149.

*****Hanninen, G. E. (1989). *The effects of the Hilda Taba teaching strategies on critical and creative thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8918535)

*****Harbeck, C. A. (1973). *Experimental study of the effect of two proof formats in high school geometry on critical thinking and selected student attitudes* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. AAT 7302013)

*****Harrell, M. (2011). Argument diagramming and critical thinking in introductory philosophy. *Higher Education Research & Development*, *30*(3), 371-385. Retrieved from http://www.hss.cmu.edu/philosophy/harrell/HarrellHERDSubNotReview.pdf

*****Harty, H., Woods, C. J., Johnson, D. A., & Pifer, S. (1986). *Academic challenge*

program: evaluation report. Indianapolis, IN.: Indiana State Dept. of Education.
*Hash, R. J. (1974). *The effects of a strategy of structured overviews, levels guides and vocabulary exercises on student achievement, reading comprehension, critical thinking and attitudes of junior high school classes in social studies* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7429239)
*Henderson, J. S. (2002). *Critical to your health: using the theory of planned behavior to teach critical thinking skills and their effect on health behaviors* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3053289)
*Hendrix, N. H. (1995). *Improving critical thinking and reading achievement in community college students.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9500238)
*Heraty, N., & Morley, M. J. (2000). The application of the structure of intellect programme: A manufacturing facility experiment. *Journal of Managerial Psychology, 15*(7), 691-715. http://dx.doi.org/10.1108/02683940010378063
*Hernstein, R., Nickerson, R., Sanchez, M., & Swets, J. (1986). Teaching thinking skills. *American Psychologist, 41*, 1279-1289. http://dx.doi.org/10.1037/0003-066X.41.11.1279
*Hesterberg, L. J. (2005). *Evaluation of a problem-based learning practice course: Do self-efficacy, critical thinking, and assessment skills improve?* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3162941)
*Hill, N. J. (1997). *Impact of a critical thinking continuing education program on the critical thinking skills of practicing nurses* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9737644)
*Hinton, R. S. (1980). *A study of the effects of utilizing local history on retention of historical information, critical thinking skills, and attitudes in teaching United States history.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8010670)
*Horn, C., & Bruning, R. (2004). The role of online discussion in promoting critical thinking using: Case studies with pre-service teachers. In L. Cantoni & C. McLoughlin (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2004* (pp. 3088-3094). Chesapeake, VA: AACE. Retrieved from http://www.editlib.org/p/12459
*Housen, A. C. (2001-2002). Aesthetic thought, critical thinking and transfer. *Arts and Learning Research Journal, 18*(1), 99-132.
*Houtveen, A. A. M., & Grift, W. J. C. M. v. d. (2007). Effects of metacognitive strategy instruction and instruction time on reading comprehension. *School Effectiveness and School Improvement, 18,* 173-190. http://dx.doi.org/10.1080/09243450601058717
*Howard, V. M. (2007). *A comparison of educational strategies for the acquisition of medical-surgical nursing knowledge and critical thinking skills: Human patient simulator vs. the interactive case study approach* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3270096)
*Hsu, L. L. (2004). Developing concept maps from problem-based learning scenario discussions. *Journal of Advanced Nursing, 48*, 510-518.

http://dx.doi.org/10.1111/j.1365-2648.2004.03233.x

*****Hunt, C. A. (2003). *Promoting critical thinking: an analysis of instructional techniques with undergraduates in a university setting* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3076096)

*****Hur, S.-J. (2001). *Effects of anchored instruction on the critical-thinking skills of students with and without mild disabilities* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3037504)

*****Inlow, F. H., & Chovan, W. (1993). Another search for the effects of teaching thinking and problem solving skills on college students' performance. *Journal of Instructional Psychology, 20*, 215-223.

*****Janiszewski Goodin, H. I. (2005). *The use of deliberative discussion as a teaching strategy to enhance the critical thinking abilities of freshman nursing students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3176411)

*****Johnson, A. P. (1997). *The effect of general thinking skills instruction on the thinking and writing performance of fifth and sixth grade students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9719668)

*****Jones, C. (1970). *The effects of participation in discussion of reasoning in arguments on critical thinking ability* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7013719)

*****Jones, H. E. (1993). *The effect of direct instruction of thinking skills in elementary social studies on the development of thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9322839)

*****Jones, J. T. (1975). *An experimental study of four interdisciplinary approaches to promoting critical thinking skills and personal development in the college classroom* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7503500)

*****Joung, S. (2004). *The effects of high-structure cooperative versus low-structure collaborative design on online debate in terms of decision-making, critical thinking, and interaction pattern* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3109292)

*****Joyce, J. (1971). *A comparison of two methods of teaching life career planning to junior high school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7204087)

*****Kaberman, Z., & Dori, Y. J. (2009). Question posing, inquiry, and modeling skills of chemistry students in the case-based computerized laboratory environment. *International Journal of Science and Mathematics Education, 7*, 597-625. http://dx.doi.org/10.1007/s10763-007-9118-3

*****Kaberman, Z., & Dori, Y. J. (2009). Metacognition in chemical education: question posing in the case-based computerized learning environment. *Instructional Science, 37*, 403-436. http://dx.doi.org/10.1007/s11251-008-9054-9

*****Kaplan, I. S. (1997). *An assessment of the infusion of critical thinking skills into content instruction* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9737548)

*****Keller, R. (1994). *Effects of an instructional program on critical thinking and clinical decision-making skills of associate degree nursing students (nursing education)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9404905)

*Kemp, S. G., & Sadoski, M. (1991). The effects of instruction in forming generalizations on high school students' critical thinking in world history. *Reading Research and Instruction, 31*(1), 33-42.

*Khoiny, F. E. (1996). *The effectiveness of problem-based learning in nurse practitioner education* (Doctoral dissertations). Available from ProQuest Dissertations and Theses database. (UMI No. 9614036)

*Kirschbaum, H. (1979). *The effects of presenting general semantics lessons in five media upon the development of critical thinking in junior high school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7925272)

*Klauer, K. J. (1996). Teaching inductive reasoning: some theory and three experimental studies. *Learning and Instruction, 6*, 37-57. http://dx.doi.org/10.1016/S0959-4752(96)80003-X

*Knight, D. H. (1981). *The effects of transfer materials on the critical thinking abilities of second-year algebra students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8107220)

*Knoff, E., Rugh, J. D., & Littlefield, J. H. (1990). Comparing learning logs and note summaries in a dental research methods course. *Journal of Dental Education, 54*, 724-729.

*Knotts, J. D. (1998). *Differential effectiveness of two instructional methods for developing higher-order question writing skills of preservice special educators* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9808829)

*Koehler, C. F. (1994). *The effectiveness of deconstructive teaching strategies on students' critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9409422)

*Koray, O., & Koksal, M. S. (2009). The effect of creative and critical thinking based laboratory applications on creative and logical thinking abilities of prospective teachers. *Asia-Pacific Forum on Science Learning and Teaching, 10*(1). Retrieved from http://www.ied.edu.hk/apfslt/v10_issue1/koksal/

*Korkmaz, O., & Karakus, U. (2009). The impact of blended learning model on student attitudes towards geography course and their critical thinking dispositions and levels. *Turkish Online Journal of Educational Technology, 8*(4), 51-63.

*Kowalski, S. D., & Louis, M. A. (2000). Using computer technology to teach critical thinking. *Nurse Educator, 25*, 210-212.

*Kramarski, B., & Zoldan, S. (2008). Using errors as springboards for enhancing mathematical reasoning with three metacognitive approaches. *Journal of Educational Research, 102*, 137-151.

*Krank, H. M. (1994). *Do cognitive styles and learning environment predict critical thinking performance for pre-service teachers?* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9405077)

*Kumta, S., Tsang, P., Hung, L., & Cheng, J. (2003). Fostering critical thinking skills through a web-based tutorial programme for final year medical students – a randomized controlled study. *Journal of Educational Multimedia and Hypermedia, 12*, 267-273.

*Langer, P., & Chiszar, D. (1993). Assessment of critical thinking courses. *Perceptual* *and Motor Skills, 77*, 970.

*Lee, C., Tsai, H., & Lin, C. (2003). A Web-based Instructional Design for Teaching Critical Thinking Ability---Using Current Issues As An Example. In C. Crawford et al. (Ed.), *Proceedings of the Society for Information Technology and Teacher Education International Conference 2003* (pp. 3281-3284). Chesapeake, VA: AACE.

*Lee, K. (2005). Enhancing critical thinking in online learning. *Academic Exchange Quarterly, 9*(4), 43-48.

*Lee, K. (2007). Online collaborative case study learning. *Journal of College Reading and Learning, 37*(2), 82-100.

*Lee, M. O. C., & Thompson, A. (1997). Guided instruction in logo programming and the development of cognitive monitoring strategies among college students. *Journal of Educational Computing Research, 16*, 125-144. http://dx.doi.org/10.2190/PW3F-HLFD-1NNJ-H77Q

*Leonard, M. (2007). *Teaching critical thinking in a community-college art-history learning environment: A comparison of implicit instruction versus a combination of implicit and explicit instruction* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3280031)

*Lewis, H. (1962). *An experiment in developing critical thinking through the teaching of plane demonstrative geometry* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 0001926)

*Lin, C.-C., Lee, C.-J., & Tsai, H.-L. (2003). A web-based instructional design for teaching critical thinking ability using current issues as an example. In C. Crawford et al. (Ed.), *Proceedings of the Society for Information Technology and Teacher Education International Conference* (pp. 3281-3284). Chesapeake, VA: AACE. Retrieved from http://www.editlib.org/p/18693

*Lindsay, C. W. (1974). *A study of the effects of three methods of teaching high school chemistry upon achievement in chemistry, critical thinking abilities, and scientific interest* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7411454)

*Little, C. A., Feng, A. X., VanTassel, B. J., Rogers, K. B., & Avery, L. D. (2007). A study of curriculum effectiveness in social studies. *Gifted Child Quarterly, 51*(3), 272-284.

*Liu, C.-C. (2006). *Connecting the dots between critical thinking skills and dispositions for middle school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3243081)

*Liu, M., & Pedersen, S. (1998, June). *The effect of being hypermedia designers on elementary school students' motivation and learning of design knowledge*. Paper presented at the ED-MEDIA/ED-TELECOM 98 World Conference on Educational Multimedia and Hypermedia & World Conference on Educational Telecommunications, Freiburg, Germany. (ERIC Document Reproduction Service No. ED428689)

*Livingston, H. (1964). *The effect of instruction in general semantics on the critical reading ability of 10th grade students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6506581)

*Lohse, B., Nitzke, S., & Ney, D. M. (2003). Introducing a problem-based unit into a lifespan nutrition class using a randomized design produces equivocal outcomes. *Journal of the American Dietetic Association, 103*, 1020-1025.

http://dx.doi.org/10.1016/S0002-8223(03)00478-4
*Loiacono, C. M. (2000). *Promoting freshman college student development using cognitive developmental theory presented in a deliberate psychological education-based freshman orientation program* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9974946)
*Loveless, E. M. (1970). *Developing critical thinking skills in students in eighth grade English classes* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7013720)
*Luckett, R. A. (1991). *The effect of the use of algorithmization on critical thinking skills of upper-division college students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9127000)
*Lumpkin, C. R. (1991). *Effects of teaching critical thinking skills on the critical thinking ability, achievement, and retention of social studies content by fifth and sixth-graders (fifth- graders)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9109097)
*Lyons, E. B. (2006). *Examining the effects of problem-based learning on the critical thinking skills of associate degree nursing students in a southeastern community college* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3227670)
*Lyons, E. M. (2008). Examining the effects of problem-based learning and NCLEX-RN scores on the critical thinking skills of associate degree nursing students in a Southeastern Community College. *International Journal of Nursing Education Scholarship, 5*, http://dx.doi.org/10.2202/1548-923X.1524
*Macklem, R. E. (1974). *The effects of a specific teacher inservice program in BSCS biology upon teacher and student verbal behaviors and selected student learning outcomes.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7413000)
*MacPhail Wilcox, B., Dreyden, J., & Eason, E. (1990). An investigation of paideia program effects on students' critical thinking skills. *Educational Considerations, 17*(2), 61-67.
*Maina, M. P. (1997). *Effects of divergent production style h initiative game activities on fifth-grade students use of critical thinking skills in physical education* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9725010)
*Mandell, A. (1967). *A study of the effects of laboratory experiences upon critical thinking ability in a college biology course* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6701026)
*Manning, W. J. H. (1999). *The relationship between critical thinking and attitudes toward reading of the community college student enrolled in a critical reading course at Roane State Community College* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9903929)
*Martin, C. L. (2003). *The effect of training on students' critical thinking skills in computer conferences over time* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3061096)
*Massey, S. L. (2004). *The effects of cooperative learning versus traditional classroom instruction on cognitive achievement, critical thinking, and attitudes toward learning in teams in a physician assistant program* (Doctoral dissertation).

Available from ProQuest Dissertations and Theses database. (UMI No. 3098166)

*Mazer, J. P., Hunt, S. K., & Kuznekoff, J. H. (2007). Revising general education: Assessing a critical thinking instructional model in the basic communication course. *Journal of General Education 56*, 173-199

*Mcaulay, D. L. (1988). *Type ii enrichment by gifted/talented teachers: the effects of inservicing critical thinking on third-grade teachers and students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8728889)

*McBride, R., & Bonnette, R. (1995). Teacher and at-risk students' cognitions during open-ended activities: structuring the learning environment for critical thinking. *Teaching and Teacher Education, 11*, 373-388. http://dx.doi.org/10.1016/0742-051X(94)00040-D

*McCarthy-Tucker, S. N. (1995). *Teaching reality-based formal logic to adolescents to improve critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9530227)

*McCarthy-Tucker, S. N. (1998). Teaching logic to adolescents to improve thinking skills. *Korean Journal of Thinking and Problem Solving, 8*(1), 45-66.

*McCormick, G. J. (1989). *The effects of using print media in teaching critical thinking skills to adult students.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8901471)

*McCrink, C. L. S. (1999). *The role of innovative teaching methodology and learning styles on critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9905019)

*Mcgregor, G. D. J. (2002). *Creative thinking instruction for a college study skills program: A case study* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3027933)

*McKee, M. L. (2006). *Overcoming obstacles in teaching college students how to think critically: The effects of an instructional intervention* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3222188)

*McKeon, L. M., Norris, T., Cardell, B., & Britt, T. (2009). Developing patient-centered care competencies among prelicensure nursing students using simulation. *Journal of Nursing Education, 48*(12), 711-715. http://dx.doi.org/10.3928/01484834-20091113-06

*Mcmurray, M. A. (1989). *Biology analytical problem solving interventions and field independence as predictors of critical thinking changes* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8902756)

*Meiss, G. T., & Bates, G. W. (1984, May). *Cognitive and attitudinal effects of reasoning message strategies*. Paper presented at the 34th annual meeting of the International Communication Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED246519)

*Miri, B., Ben-Chaim, D., & Zoller, U. (2007). Purposely teaching for the promotion of higher-order thinking skills: A case of critical thinking. *Research in Science Education, 37*, 353-369. http://dx.doi.org/10.1007/s11165-006-9029-2

*Montean, J. J. (1960). *An experimental study of the use of discussion groups in general chemistry and general science as a means of increasing group growth in critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 5906311)

*Montuori, L. A., & Kimmel, E. B. (1994, August). *Teaching conceptual complexity to adults using an in-basket instructional design.* Paper presented at the annual meeting of the American Psychological Association, Los Angeles, CA. (ERIC Document Reproduction Service No. ED388768)

*Moreyra, A. A. (1992). *The role of thinking frames in developing teachers' critical thinking skills and dispositions* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9136492)

*Neimark, E. D. (1984, August). *A cognitive style change approach to the modification of thinking in college students*. Paper presented at the Conference on Thinking, Cambridge, MA. (ERIC Document Reproduction Service No. ED261301)

*Nielsen, M. E. (1985). *Evaluation of a rural gifted program: Assessment of attitudes, self-concepts, and critical thinking skills of high-ability students in grades 3 through 12* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8500417)

*Nieto, A. M., & Saiz, C. (2008). Evaluation of Halpem's "structural component" for improving critical thinking. *The Spanish journal of psychology, 11*, 266-274. Retrieved from http://www.ucm.es/info/Psi/docs/journal/v11_n1_2008/art266.pdf

*Noland, R. L. (2002). *The effectiveness of the ropes challenge course on the enhancement of critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3055822)

*Overton, J. C. (1994). *An investigation of the effects of thinking skills instruction on academic achievement and the development of critical and creative thinking skills of second-, fourth-, and sixth-grade students (second-grade, fourth- grade)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9417153)

*Pachtman, A. B. (1977). *The effects of a reading and language arts program on the critical thinking and critical reading of the first-year law students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7724569)

*Parker, M. J. (1999). *The effects of a shared, intranet science learning environment on the academic behaviors of problem-solving and metacognitive reflection* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9912894)

*Parkinson, M. G., & Ekachai, D. (2002). The Socratic method in the introductory PR course: An alternative pedagogy. *Public Relations Review, 28*, 167-174. http://dx.doi.org/10.1016/S0363-8111(02)00123-6

*Pellegrino, A. M. (2007). *The manifestation of critical thinking and metacognition in secondary American history students through the implementation of lesson plans and activities consistent with historical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3282653)

*Penningroth, S. L., Despain, L. H., & Gray, M. J. (2007). Topical articles: A course designed to improve psychological critical thinking. *Teaching of Psychology, 34*, 153-157. http://dx.doi.org/10.1080/00986280701498509

*Perkins, H. W. (1985). *The effects of microcomputer use on the critical thinking skills of middle school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8506714)

*Pfeffer, A. C. (1977). *The effect of an independent study process on the critical* *thinking ability of seventh grade middle school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7713522)

*Phelps, P. H. (1988). *The effects of participation in reflective teaching on preservice teachers' critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8725303)

*Pickens, C. L. (2007). *Concept mapping: Methods to improve critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3279742)

*Pierce, W., Lemke, E., & Smith, R. (1988, April). *Effects of a secondary curriculum for moral development and critical thinking.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED296940)

*Pierleoni, R. G. (1968). *The affective measurement of a programmed unit on American political attitudes and positions* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6815879)

*Pisano, R. C. (1980). *The effectiveness of an intervention study in critical thinking skills designed to improve written composition in eleventh and twelfth graders* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8014250)

*Poel, R. H. (1971). *Critical thinking as related to PSSC and non-PSSC physics programs* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7103944)

*Quitadamo, I. J. (2003). *Critical thinking in higher education: The influence of teaching styles and peer collaboration on science and math learning* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3069650)

*Quitadamo, I. J., Faiola, C. L., Johnson, J. E., & Kurtz, M. J. (2008). Community-based inquiry improves critical thinking in general education biology. *CBE - Life Sciences Education, 7*, 327-337. http://dx.doi.org/10.1187/cbe.07-11-0097

*Quitadamo, I. J., & Kurtz, M. J. (2007). Learning to improve: Using writing to increase critical thinking performance in general education biology. *CBE Life Sciences Education, 6*, 140-154. http://dx.doi.org/10.1187/cbe.06-11-0203

*Ray, C. L. (1979). *A comparative laboratory study of the effects of lower level and higher level questions on students' abstract reasoning and critical thinking in two non-directive high school chemistry classrooms* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7926803)

*Reed, J. H. (1999). *Effect of a model for critical thinking on student achievement in primary source document analysis and interpretation, argumentative reasoning, critical thinking dispositions, and history content in a community college history course* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9911510)

*Reed, J. H., & Kromrey, J. D. (2001). Teaching critical thinking in a community college history course: Empirical evidence from infusing Paul's model. *College Student Journal, 35*, 201-215.

*Reid, E. C. (1982). *Training higher-level cognitive skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8215823)

*Renaud, R. D. (2002). *The effect of higher-order questions on critical thinking skills*

(Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. NQ67183)

*Resnick, L. B., Bill, V., & Lesgold, S. (1992). Developing thinking abilities in arithmetic class. In A. Demetriou, M. Shayer & A. Efklides (Eds.), *Neo-Piagetian Theories of Cognitive Development* (pp. 210-230). London: Routledge. Retrieved from

*Rickert, R. K. (1967). Developing critical thinking. *Science Education, 51*, 24-27. http://dx.doi.org/10.1002/sce.3730510106

*Riley, V. M. (1992). *Teachers' questioning for improvement of critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9135488)

*Rivers, R. H., & Vockell, E. (1987). Computer simulations to stimulate scientific problem solving. *Journal of Research in Science Teaching, 24*, 403-415. http://dx.doi.org/10.1002/tea.3660240504

*Robinson, T. Y. (2005). *A study of the effectiveness of environmental education curricula in promoting middle school students' critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3195340)

*Rogers, G., & McDonald, L. (1992). Thinking critically - an approach to field instructor training. *Journal of Social Work Education, 28*, 166-177.

*Rose, M. M. (1998). *Critical thinking skills instruction for postsecondary students with and without learning disabilities: The effectiveness of icons as part of a literature curriculum* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9806188)

*Ross, G. A., & Semb, G. (1981). Philosophy can teach critical thinking skills. *Teaching Philosophy, 4*, 111-122.

*Rothert, L. F. (1969). *An analysis of changes in critical thinking ability, open-mindedness, and farm policy opinions of participants in the Kellogg farmers study program* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6920923)

*Rothstein, A. (1962). *An experiment in developing critical thinking through the teaching of American history* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6003758)

*Ruff, L. G. (2005). *The development of critical thinking skills and dispositions in first-year college students: Infusing critical thinking instruction into a first-year transitions course* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3201999)

*Ryan, L. G. (1996). *Critical thinking in social work practice: a quasi- experimental investigation* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9622836)

*Sanders, J. A., Wiseman, R. L., & Gass, R. H. (1994). Does teaching argumentation facilitate critical thinking? *Communication Reports, 7*, 27-35. http://dx.doi.org/10.1080/08934219409367580

*Sasscer, J. C. (1973). *The development, implementation, and evaluation of a modularized, student-centered general biology curriculum at the college level* (Unpublished doctoral dissertation). The Catholic University of America.

*Schroeder, J. M. (2007). *A study of improving critical thinking skills with multiple choice tests and first semester associate degree nursing students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3278154)

*Schuler, G. (1974, April). *The effectiveness of the productive thinking program.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois. (ERIC Document Reproduction Service No. ED103479)

*Schulhauser, C. E. (1991). *The effects of literary discussion groups on students' critical thinking ability and attitude toward reading* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9131095)

*Schulze, F. R. (1998). *The effects of instructional strategies on levels of abstract thought* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9815686)

*Sendag, S., & Odabasi, H. F. (2009). Effects of an online problem based learning course on content knowledge acquisition and critical thinking skills. *Computers & Education, 53*, 132-141. http://dx.doi.org/10.1016/j.compedu.2009.01.008

*Shamir, A., Zion, M., & Spector_Levi, O. (2008). Peer tutoring, metacognitive processes and multimedia problem-based learning: The effect of mediation training on critical thinking. *Journal of Science Education and Technology, 17*, 384-398. http://dx.doi.org/10.1007/s10956-008-9108-4

*Shearer, J. E. (1981). *A comparison of individualized and group-oriented reading methods related to critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8106340)

*Shepherd, N. G. (1998). *The probe method: A problem-based learning model's affect on critical thinking skills of fourth and fifth grade social studies students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9826025)

*Shin, S.-R. (2003). *The effects of a metacognitive art criticism teaching strategy that incorporates computer technology on critical thinking skill and art critiquing ability* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3072012)

*Sidney, A. H. (1989). *The effects of the inquiry method of teaching science on critical thinking skills, achievement and attitude toward science* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8917080)

*Slaughter, D. S., Brown, D. S., Gardner, D. L., & Perritt, L. J. (1989). Improving physical therapy students' clinical problem-solving skills: an analytical questioning model. *Physical Therapy*, *69*(6), 441-447.

*Slavik, S. J. (1996). *An examination of the effects of selected disciplinary art teaching strategies on the cognitive development of selected sixth grade students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9544331)

*Smith, B. (1974). *Critically reading for propaganda techniques in grade six.* (Unpublished master's thesis). Rutgers University, New Brunswick, NJ.

*Solon, T. (2001). Improving critical thinking in an introductory psychology course. *Michigan Community College Journal: Research and Practice*, *7*(2), 73-80.

*Solon, T. (2007). Generic critical thinking infusion and course content learning in introductory psychology. *Journal of Instructional Psychology, 34*, 95-109.

*Sondel, H. (2009). *The effects of curricular programs on aspects of critical thinking as applied to writing* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3364889)

*Stambaugh, T. (2007). *Effects of the Jacob's Ladder Reading Comprehension Program on reading comprehension and critical thinking skills of third, fourth, and fifth grade students in rural schools* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3254406)

*Starr, R. J. (1972). Structured oral inquiry improves thinking. *American Biology Teacher, 34*, 408-409.

*Stoiber, K. C. (1991). The effect of technical and reflective preservice instruction on pedagogical reasoning and problem solving. *Journal of Teacher Education, 42*, 131-139. http://dx.doi.org/10.1177/002248719104200206

*Suksringarm, P. (1976). *An experimental study comparing the effects of BSCS and traditional biology on achievement, understanding of science, critical thinking ability, and attitude toward science of the first year students at the Sakon Nakorn Teachers' College, Thailand* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7624805)

*Sullivan-Mann, J., Perron, C. A., & Fellner, A. N. (2009). The effects of simulation on nursing students' critical thinking scores: A quantitative study. *Newborn and Infant Nursing Reviews, 9*, 111-116. http://dx.doi.org/10.1053/j.nainr.2009.03.006

*Sungur, S., & Tekkaya, C. (2006). Effects of problem-based learning and traditional instruction on self-regulated learning. *Journal of Educational Research, 99*, 307-317. http://dx.doi.org/10.3200/JOER.99.5.307-320

*Sykes, M. S. (1991). *The effects of instructional interventions for at-risk secondary students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9119219)

*Szabo, Z., & Schwartz, J. (2009). Learning methods for teacher education: Blackboard discussions produce deep learning. In G. Siemens & C. Fulford (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009* (pp. 2323-2334). Chesapeake, VA: AACE. Retrieved from

*Tabbada, E. V. (1988). *Exploring the development of critical thinking in home economics* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8803216)

*Tarkington, S. A. (1989). *Improving critical thinking skills using Paideia seminars in a seventh grade literature curriculum* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8912761)

*Teagle, T. E. (1986). *The Socratic method of teaching: Its effect on the development of critical thinking skills of upper grade elementary school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8617404)

*Teele, L. (2006). *The impact of integrated study skills and critical thinking on student achievement* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3226249)

*Tell, M. P. (1989). *The impact of a thinking skills curriculum on selected cognitive processes of intermediate grade students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8818721)

*Thomas, J., Coppola, J. F., & Thomas, B. A. (2002). Smart e-classrooms, traditional classrooms and critical thinking. In G. Richards (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2002* (pp. 2288-2291). Chesapeake, VA: AACE. Retrieved from http://www.editlib.org/p/9750

*Thomas, J. W. (1975, March-April). *The effects of instruction on the critical thinking abilities of middle-school-age children.* Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C. (ERIC Document Reproduction Service No. ED105420)

*Thompson, J. (2009). *To question or not to question: The effects of two teaching approaches on students' thinking dispositions, critical thinking skills, and course grades in a critical thinking course* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3355859)

*Thompson, S. S. (1995). *Techniques for assessing general education outcomes: the natural science core program at the University of Denver (Colorado)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9532955)

*Thomson, S. (2009). *Effects of Primary Education Thinking Skills (TM) on achievement: Mixed-method evaluative study* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3401313)

*Tilson, E. R. (1987). *The effect of computer enhanced skill training in critical thinking and cognitive monitoring on learning and development of critical thinking in an undergraduate family nursing course (Georgia)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8706902)

*Tiwari, A., Lai, P., So, M., & Yuen, K. (2006). A comparison of the effects of problem-based learning and lecturing on the development of students' critical thinking. *Medical Education, 40*, 547-554. http://dx.doi.org/10.1111/j.1365-2929.2006.02481.x

*Topping, K. J., & Trickey, S. (2007). Collaborative philosophical inquiry for schoolchildren: Cognitive gains at 2-year follow-up. *British Journal of Educational Psychology, 77*, 787-796. http://dx.doi.org/10.1348/000709907X193032

*Unan, V. S. S. (1979). *A comparison of the effects of BSCS and traditional biological science on achievement, critical thinking ability, and attitude toward science, scientists, and scientific careers of female slow learners in an Egyptian preparatory school* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7902661)

*VanderStoep, S. W., & Shaughnessy, J. J. (1999). Taking a course in research methods improves reasoning about real-life events. In M. Ware & C. L. Brewer (Eds.), *Handbook for teaching statistics and research methods* (2 ed., pp. 242-244). Mahwah, NJ: Lawrence Erlbaum.

*Vanlengen, C. A. (1989). *Effect of computer programming instruction on the problem solving ability of college-level introductory computer students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8907819)

*VanTassel-Baska, J., Bracken, B., Feng, A., & Brown, E. (2009). A longitudinal study of enhancing critical thinking and reading comprehension in title I classrooms. *Journal for the Education of the Gifted, 33*(1), 7-37.

*Vasu, E. S., & Tyler, D. K. (1997). A comparison of the critical thinking skills and spatial ability of fifth grade children using simulation software or logo. *Journal of Computing in Childhood Education, 8,* 345-363.

*Velde, B. P., Wittman, P. P., & Vos, P. (2006). Development of critical thinking in occupational therapy students. *Occupational Therapy International*, *13*, 49-60. http://dx.doi.org/10.1002/oti.20

*Waldron, J. M. (1992). *The effects of two instructional strategies for critical thinking-writing instruction on high ability middle school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9231949)

*Wallace, S. P. (1993). *A study of argumentative/persuasive writing related to a model of critical thinking in grades nine and eleven* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9239095)

*Wallen, A., Reid, T., & Haubrick, V. (1963). The outcomes for critical thinking. *Journal of Educational Research*, *56*, 529-534.

*Wang, Y. (2009). Using discussion forum to enhance students' critical thinking ability through the Blackboard learning system. In T. Bastiaens et al. (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009* (pp. 3243-3252). Chesapeake, VA: AACE.

*Weast, D. (1996). Alternative teaching strategies: The case for critical thinking. *Teaching Sociology*, *24*, 189-194.

*Wells, L. A. (2002). Philosophy and psychiatry - A new curriculum for child and adolescent psychiatry. *Academic Psychiatry, 26*, 257-261. Retrieved from http://ap.psychiatryonline.org/cgi/reprint/26/4/257

*Whalen, S. (1991). Intercollegiate debate as a cocurricular activity: Effects on critical thinking. In D. Parson (Ed.), *Argument in controversy: Proceedings of the 7th SCA/AFA Conference on Argumentation* (pp. 391-397). Annandale, VA: Speech Communication Association.

*Whitten, M. M. (1971). *Changes in understanding of the nature of science and in critical thinking resulting from two different physical science courses* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7122549)

*Whyte, D. T. (1999). *The effect of an educational unit on the critical thinking skills of social work students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9841778)

*Williams, R. L., Oliver, R., & Stockdale, S. (2004). Psychological versus generic critical thinking as predictors and outcome measures in a large undergraduate human development course. *Journal of General Education*, *53*, 37-58.

*Winocur, S. M. (1981). *The impact of a program of critical thinking on reading comprehension remediation and critical thinking of middle and high school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8116378)

*Wolf, W., King, M. L., & Huck, C. S. (1968). Teaching critical reading to elementary school children. *Reading Research Quarterly*, *3*, 435-498.

*Wright, A. (1991). *Critical thinking: Definition, measurement and development in vocational students in further education* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. DX93650)

*Wu, W., & Yang, Y. C. (2009). Using a Multimedia Storytelling to Improve

Students' Learning Performance. In G. Siemens & C. Fulford (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009* (pp. 3159-3166). Chesapeake, VA: AACE. Retrieved from

*Xie, Y., Ke, F., & Sharma, P. (2008). The effect of peer feedback for blogging on college students' reflective learning processes. *Internet & Higher Education 11*, 18-25 http://dx.doi.org/10.1016/j.iheduc.2007.11.001

*Yalcin, B. M., Karahan, T. F., Karadenizli, D., & Sahin, E. M. (2006). Short-term effects of problem-based learning curriculum on students' self-directed skills development. *Croatian Medical Journal*, *47*(3), 491-498. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2080434/pdf/CroatMedJ_47_0491.pdf

*Yang, S. C., & Chung, T.-Y. (2009). Experimental study of teaching critical thinking in civic education in Taiwanese junior high school. *British Journal of Educational Psychology*, *79*, 29-55. http://dx.doi.org/10.1348/000709907X238771

*Yang, Y., & Chou, H. (2005). E-Learning: Learning How to Think through Interaction via Computer Conferencing. In P. Kommers & G. Richards (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2005* (pp. 886-890). Chesapeake, VA: AACE. Retrieved from

*Yang, Y.-T. C. (2004). *Use of structured web-based bulletin board discussions with Socratic questioning to enhance students' critical thinking skills in distance education* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3105051)

*Yang, Y.-T. C., & Chou, H.-A. (2008). Beyond critical thinking skills: Investigating the relationship between critical thinking skills and dispositions through different online instructional strategies. *British Journal of Educational Technology 39*, 666-684 http://dx.doi.org/10.1111/j.1467-8535.2007.00767.x

*Yang, Y.-T. C., Newby, T., & Bill, R. (2008). Facilitating interactions through structured web-based bulletin boards: A quasi-experimental study on promoting learners' critical thinking skills. *Computers & Education 50*, 1572-1585 http://dx.doi.org/10.1016/j.compedu.2007.04.006

*Yang, Y. C., & Li, L. (2009). Training pre-service teachers in how to teach critical thinking to middle school students. In T. Bastiaens et al. (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009* (pp. 2060-2069). Chesapeake, VA: AACE. Retrieved from

*Yarema, C. H. (1996). *Effects of use of a context approach to calculus on students' critical thinking abilities, content acquisition, and attitudes toward mathematics* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9600133)

*Yeh, Y. C. (2009). Integrating e-learning into the direct-instruction model to enhance the effectiveness of critical-thinking instruction. *Instructional Science*, *37*, 185-203. http://dx.doi.org/10.1007/s11251-007-9048-z

*Yuan, H., Kunaviktikul, W., Klunklin, A., & Williams, B. A. (2008). Improvement of nursing students' critical thinking skills through problem-based learning in the People's Republic of China: A quasi-experimental study. *Nursing & Health Sciences 10*, 70-76 http://dx.doi.org/10.1111/j.1442-2018.2007.00373.x

*Yuill, N. C. W. (1992). *The effects of computerized problem-solving simulations on critical thinking skills of baccalaureate nursing students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9203063)

*Zohar, A., & Tamir, P. (1993). Incorporating critical thinking into a regular high school biology curriculum. *School Science and Mathematics, 93*, 136-140.
*Zohar, A., Weinberger, Y., & Tamir, P. (1994). The effect of the biology critical thinking project on the development of critical thinking. *Journal of Research in Science Teaching, 31*, 183-196. http://dx.doi.org/10.1002/tea.3660310208

5. Pedagogical advice

- Cook, J. (2017). How to Effectively Debunk Myths About Aging and Other Misconceptions. *Public Policy & Aging Report*, 27(1), 13-17.
- Maurage, P., Heeren, A., & Pesenti, M. (2013). Does chocolate consumption really boost Nobel award chances? The peril of over-interpreting correlations in health studies. *The Journal of Nutrition*, 143(6), 931-933.
- Messerli FH, *Chocolate consumption, cognitive function, and Nobel laureates*. N Engl J Med. 2012; 367 :1562–4.
- Matute, H., Blanco, F., Yarritu, I., Díaz-Lago, M., Vadillo, M. A., & Barberia, I. (2015). Illusions of causality: how they bias our everyday thinking and how they could be reduced. *Frontiers in Psychology*, *6*, 888.